ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

BIWI Computer Vision

Motion Correction and IVIM Fitting for in-utero Fetal Diffusion-weighted Magnetic Resonance Images

Master's Thesis

Lin Zhang

Department of Information Technology and Electrical Engineering
ETH Zurich

Supervisors

Dr. Valery Vishnevskiy
PD Dr. med. Dr. phil. András Jakab
Prof. Dr. Orcun Göksel

November 4, 2018

Abstract

Intravoxel incoherent motion magnetic resonance imaging (IVIM-MRI) allows contrast-agent free in vivo perfusion quantification in developing human fetus. However, clinical translation of prenatal IVIM imaging is limited by poor estimation accuracy from low signal-to-noise (SNR) ratio due to strong diffusion encoding, spatial misalignment and dephasing artefacts induced by random fetal motion. To address these issues, we define an implicit signal acquisition model considering non-Gaussian noise and signal dephasing. An artificial neural network is proposed to learn entire posterior of IVIM parameters with this model. This allows uncertainty quantification of the inferred parameters, which we validate with true posterior approximated by Bayesian sampling. Together with a registration-based motion correction pipeline, the proposed method is evaluated on in vivo fetal MR images. Compared to the conventional least squares (LSQ) approach, this approach achieves higher estimation accuracy on synthetic data and increases repeatability of parameter estimation in placenta for the in vivo cases.

Acknowledgements

I would like to express my profound gratitude to my supervisors Dr. Valery Vishnevskiy, PD Dr. András Jakab and Prof. Dr. Orcun Göksel for continuous motivation, guidance and suggestion during my master thesis. I am very thankful for the support in developing a conference paper from the thesis. Their experience, feedback and insight greatly assisted the research. I very appreciate this experience and it sets the stage for my further academic carrier.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Focus of this Work

1.1.1 Intravoxel incoherent motion model and prenatal diagnosis

Diffusion-weighted imaging (DWI) is a non-invasive method to generate image contrast based on Brownian motion in clinical magnetic resonance imaging [25]. It can provide information about microscopic details of tissue structure. In the classical DWI model, signal attenuation is explained by a mono-exponential decay. Despite of water molecule diffusion, blood perfusion in the capillaries can contribute to DWI signals as well. Therefore, the mono-exponential model might be insufficient to describe the underlying physiological phenomenon. The intravoxel incoherent motion (IVIM) model assumes a bi-exponential model to separate the diffusion of water molecules and perfusion [4, 21, 20]. The two types of translational motions take place at different temporal scales and the separation is therefore possible.
The IVIM imaging has achieved success in imaging highly perfused organs such as liver, kidney and brain [39, 23, 5]. In recent years, adapting this technique to prenatal diagnostic imaging has attracted increasing attention. Previous studies have shown that assessing the placenta functionality via perfusion quantification using IVIM imaging enables assessment of placental development, therefore makes early detection of fetal growth abnormalities possible [2, 32]. IVIM imaging has been applied on other developing organs like fetal liver, lungs, kideys and brain in [15, 16]. Their results have revealed the biological relation between IVIM parameters and microstructural development of fetal organs during gestation. Therefore, IVIM imaging can be potentially served as a useful and powerful diagnostic tool for detecting prenatal diseases. However, the applicability of IVIM imaging in clinic is currently limited by large error due to low signal-to-noise (SNR) ratio by diffusion encoding and motion artefacts induced by unpredictable fetal movement.

1.1.2 Motion Correction

Image quality is affected mainly by presence of bulk motion. Maternal and fetal motion are the two main motion sources. Maternal motion could be predictable e.g. maternal respiratory and cardiac motion. There are several clinically successful methods to reduce such physiological motion. Breath-holding or free-breathing respiratory gating [11, 22], cardiac gating [8, 19] strategies are commonly used during acquisition. However, correcting unpredictable motion such as fetal motion and maternal body movement requires more complicated approaches. Detection and correction of motion artifacts can be differentiated into two types: prospective or retrospective, summarized in [24]. Prospective methods use navigators or external tracking

devices to monitor motion during the scan [7]. Retrospective methods estimate k-space data according to the given information about subject motion [37]. These techniques require patient cooperation, external devices or longer scan time.
Motion correction based on image registration possess an alternative solution to correct spatial inconsistencies of acquired images. The images with different b-values are aligned before model fitting procedure. The conventional approach is to register each diffusion-weighted image to a non-diffusion-weighted reference image ($b = 0$ mm/s^2) using normalized mutual information [29] or cross correlation [13] as cost function. However, strong diffusion encoding gradients make diffusion weighted imaging not only sensitive to incoherent motion of water molecules, but also to subject physiological motion, leading to signal loss due to spin dephasing within the voxels. Large image contrast difference due to tissue properties and subject motion limits the registration performance. To address these issues, we propose to use group-wise registration to simultaneously align all the images to a virtual reference frame. Advantages include elimination of bias introduced by the chosen reference image and less sensitive to image contrast changes. We use the registration algorithm in [38] combined with a cost function based on principle component analysis (PCA) proposed in [14].

1.1.3 IVIM inverse mapping

The use of prenatal IVIM imaging in clinic is limited by poor IVIM model fitting performance due to low SNR of images with strong diffusion encoding gradients and motion induced dephasing artefacts, which can not be reduced by registration-based motion correction algorithms. IVIM model fitting means inferring IVIM parameters from observed signals, which is a challenging inverse mapping problem. The state-of-art fitting algorithms include maximum likelihood estimation (MLE) [27, 30] and Bayesian fitting [28, 33]. The accuracy of maximum likelihood estimation depends on image SNR because of likelihood nonlinearities. Accurate estimates require high SNR, which is not easily achievable in clinic settings [40]. MLE neglects spatial information and fits the model voxelwise. The fitted parameter maps are often relatively noisy, therefore only useful for quantification. The Bayesian fitting approaches in [28, 33] are more robust to noise, since information across a delineated region of interest (ROI) is considered in the estimatin procedure. However, the ROIs are assumed to be homogeneous and the Bayesian approaches are therefore not sensitive to tissue contrast.
Recently, a learning-based fitting algorithm was proposed in [3] to solve the IVIM inverse mapping problem in a more robust way and demonstrated for brain MRI. Compared to conventional least squares fitting, the proposed artificial neural network in [3] achieves higher estimation accuracy and provides better tissue contrast. However, the inverse mapping is not uniquely solvable due to uncertainties during the measurement, e.g. noise and signal dephasing. Therefore, point estimates are not sufficient for analyzing estimation results. The quantification of estimation uncertainty for the IVIM model was investigated in [40] using the Cramér-Rao bound, which provides the theoretical lower bound on the estimation uncertainty of an unbiased estimator. But the true estimation uncertainty can not be revealed by such approach. In this work, we propose to learn full posterior distribution approximated by mixture Gaussian. Its parameters are parametrized by a neural network trained with maximization of log-likelihood. The learning is facilitated by defining an implicit IVIM signal acquisition model based on the available in vivo fetal MR. We evaluate the proposed approach on both simulated and in vivo fetal MR data. Compared to state-of-the-art IVIM estimation methods, our method achieves higher estimation accuracy and improves intra-subject repeatability of the in vivo fetal MR inside placenta.

1.2 Thesis Organization

We begin with an outline of related literature for registration-based motion correction in quantitative magnetic resonance imaging and state-of-the-art IVIM estimation algorithms in Chapter 2. Followed by that, we introduce the prior art of IVIM inverse mapping and present our proposed inference network in Chapter 3. Chapter 4 gives a brief description of the registration algorithm used to apply on the in vivo fetal MR. The proposed fitting method is evaluated on the synthetic data and compared to a least square, Bayesian and another learning-based approach in Chapter 5. Experiments on the in vivo data are decribed in Chapter 6, including registration and IVIM inverse mapping results. Finally we close the thesis with a brief summary and future works.

Chapter 2

Related Work

Motion Correction

IVIM MRI is a quantitative magnetic resonance imaging (qMRI), aiming for quantification of tissue properties with a signal model and a series of images acquired with different acquisition parameters. Prior to model fitting procedure, the images need to be registered to correct spatial misalignment mostly due to subject motion. The main challenge is the contrast difference for the images to be registered. The conventional approaches to image registration for qMRI are based on pairwise registration with image dissimilarity metrics like mutual information [29] and cross correlation [13], which are robust against intensity changes. However, such approaches require a chosen reference image, which may bias the registration results. For IVIM imaging, common choices are non-diffusion-weighted image ($b = 0$ mm/s^2) [27] and averaged image over different b-values [15].
To eliminate the need for a reference, several registration algorithms have been proposed, which can be divided into two groups: model-based and image-based registration. Model-based registration approaches utilize the signal model to improve registration performance. An estimation framework with simultaneous image registration and model estimation has been proposed in [18] for abdominal IVIM MRI. The proposed method registers DW images, reconstruct high SNR images and estimate the IVIM parameters from the aligned images in an iterative fashion. However, such approaches rely on the assumption, that the model can fully explain the signals, which is usually not valid in practice. In comparison, image-based registration approaches align the images groupwise to a virtual reference space based on given data. A principle component analysis (PCA) based group registration was proposed in [14] as a generic registration method for qMRI. The proposed PCA-based image dissimilarity metric comes from the idea that a low-dimensional acquisition model can describe the image intensity changes. Therefore, misaligned images becomes more complex and changes the eigenvalue spectrum of the data correlation matrix.

IVIM inverse mapping

The state-of-art IVIM inverse mapping algorithms can be grouped into maximum likelihood estimation (MLE) and Bayesian estimation. In the MLE framework, least squares fitting, which is maximum likelihood estimation with Gaussian noise assumption, is widely used, including biexponential fit [30] and segmented least squares technique [27, 15, 16]. Direct fitting to the biexponential model results in unreliable estimates due to high nonlinearity of the model. The segmented least squares approach is commonly used to make the optimization more stable. In this approach, a diffusion-only mono-exponential model is first fitted to high b-values to fix the diffusion parameter and then the biexponential model is fitted to all b-values to estimate

the perfusion parameters. This approach is based on the assumption that perfusion has negligible effect at high b-values. While MLE-based methods often estimate the IVIM parameters voxelwise, thus provide noisy parameter maps, the Bayesian approach in [28, 33] fit the voxels in a region of interest simultaneously. Compared to MLE, this method is more robust to noise, since it incorporates the information across ROI into the prior distribution. This data-driven prior leads to a shrinkage effect, which means that estimates are shrunk towards the mean of the prior distribution. Therefore, the Bayesian approach can potentially reduce tissue contrast.

The authors in [3] proposed an artificial neural network to fit the combined IVIM-kurtosis model for human brain MR. Their approach, involving minimizing a sum-of-squared error function, learns the conditional expectation of parameter prediction based on a simplified signal acquisition model, which only takes noise into account. This conventional approach can lead to mode averaging, when this inverse problem could yield one-to-many mapping. To tackle this problem, learning full posterior for the prediction of continuous variables with neural network was investigated in [6, 26], where posterior is modelled as mixture Gaussian with distribution parameters inferred from an artificial neural network. Successful application has been shown with a robot kinematics problem in [6].

Chapter 3

Methods: Inverse Mapping of IVIM

3.1 Introduction

3.1.1 Intravoxel Incoherent Motion Model

Intravoxel incoherent motion imaging was proposed to model the signal acquired with diffusion MRI [20]. Diffusion refers to random displacement of water molecules due to thermal energy, which can be well described as a random walk. Molecule displacement due to free diffusion follows a Gaussian distribution. Blood microcirculation in the randomly oriented capillaries results in collective movement of blood water molecules. Both physical phenomena cause signal attenuation in the presence of diffusion sensitizing gradients [34] shown in Fig. 3.1. Stationary spins are unaffected by the paired pulse gradients, since phase

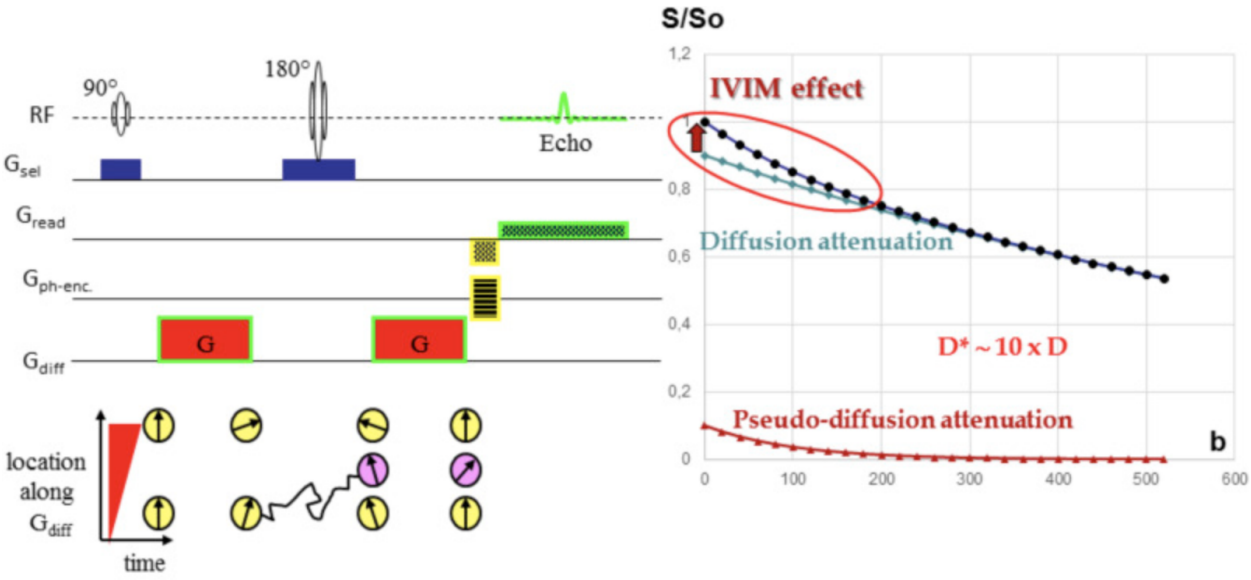


Figure 3.1: Effects of diffusion and perfusion on MRI signals. Figure courtesy [20].

accumulation by the application of the first gradient is reversed after the 180-degree pulse. On the other hand, diffusing molecules move to a different location between the paired gradients, cannot be rephased and therefore signals are lost. The degree of attenuation depends on the molecule displacement and the sensitiv-

ity of applied gradients, which is quantified as so called b-value. IVIM model aims to separate diffusion of water molecules and perfusion by describing the signal attenuation with a biexponential model [21].
The slow-moving component arising from thermal Brownian motion is quantified with the apparent diffusion coefficient D, whereas the fast-moving component governed by perfusion is described as the pseudo-diffusion coefficient D^*. The model is written as:

$$S(b;\mathbf{y}) = S_0\left(f\exp(-bD^*) + (1-f)\exp(-bD)\right), \tag{3.1}$$

where S_0 is the trace signal intensity without diffusion encoding and f is the perfusion volume fraction. $S(b)$ is the trace signal intensity acquired with the b-factor equal to b s/mm^2. Fig. 3.1 shows a graph description of the IVIM model. Pseudo diffusion appears to affect signal attenuation mainly at low b-values.
To quantify perfusion and diffusion, We have $n_p = 4$ model parameters $\mathbf{y} = \{S_0, f, D, D^*\} \in \mathbb{R}^{n_p}$ to estimate given an observed signal $\mathbf{x} \in \mathbb{R}^{n_b}$ for set of n_b b-values $\mathbf{b}$. Each element in the vector $\mathbf{x}$ represents the trace signal $S(b)$ for the b-value b. Inferring IVIM parameters $\mathbf{y}$ from an observed signal $\mathbf{x}$ is a nonlinear inverse mapping problem.

3.1.2 Prior Art

Segmented least squares approach

Given an observed signal $x \in \mathbb{R}^{n_b}$ with the b-value set $\mathbf{b}$, the inverse mapping problem can be solved in least-squares sense:

$$\hat{\mathbf{y}}_{\text{LS}} = \text{argmin}_{\mathbf{y}} \|S(\mathbf{b};\mathbf{y}) - \mathbf{x}\|_2^2. \tag{3.2}$$

Least squares method quantifies the model fit as the residuals between observed signals and values predicted by the fitted model. The estimated parameters by LSQ is equivalent to maximum likelihood estimator with Gaussian noise assumption. The segmented least square fitting (LSQ) procedure [27] is commonly used, which can stabilize the optimization procedure by avoiding some local minima. This approach is based on the assumption that perfusion has negligible effects at high b-values. In the first step, the mono-exponential diffusion model is fitted at high b-values ($b \geq 250\text{s}/\text{mm}^2$):

$$S(b;\mathbf{y}) = S_0 \cdot (1-f)\exp(-bD) = S_{int} \cdot \exp(-bD), \tag{3.3}$$

where S_{int} is called intercept and allows to eliminate S_0 in the biexponential model. In the second step, perfusion D^* and its fraction f are fitted for all b-values using fixed diffusion D and S_{int} estimates:

$$S(b;\mathbf{y}) = S_0\left(f\exp(-bD^*) + (1-f)\exp(-bD)\right) \tag{3.4}$$

$$= S_{int}\left(\frac{f}{1-f}\exp(-bD^*) + \exp(-bD)\right). \tag{3.5}$$

In practice, box constraints and an inequality are used to avoid infeasible parameter values [33]:

$$\begin{aligned} D &\leq D^* \\ 4.5\cdot 10^{-5} \leq D &\leq 1.8\cdot 10^{-2} s/\text{mm}^2 \\ 3.4\cdot 10^{-4} \leq D^* &\leq 1.0\cdot 10^{-1} s/\text{mm}^2 \\ 0.0005 \leq f &\leq 0.9995. \end{aligned} \tag{3.6}$$

Bayesian shrinkage prior inference

Due to measurement noise and other uncertainties like dephasing, this inverse problem is not uniquely solvable. Instead of point estimation, we can interpret the inverse mapping in a probabilistic way via Bayesian inference. A Bayesian estimation approach combined with data-driven priors is proposed in [28] for regional IVIM modeling, hereafter referred as BSP. The prior parameters are estimated from data in a region of interest (ROI), so that spatial information can be incorporated into inference procedure.
In the Bayesian framework, the IVIM signal model is defined as

$$x_n = S_0 \left(f \exp(-bD^*) + (1-f)\exp(-bD)\right) + \epsilon_n \tag{3.7}$$

with observed signal x_n at the n-th b-value b_n and noise term ϵ_n, which is Gaussian distributed with zero mean and variance σ_x^2. The data likelihood of observed signals $\mathbf{x} \in \mathbb{R}^{n_b}$ for set of n_b b-values $\mathbf{b}$ is written as

$$p(\mathbf{x}|f, D, D^*, S_0, \sigma_x^2) = (2\pi\sigma_x^2)^{-n_b/2} \exp\left(-\frac{1}{2\sigma_x^2}\sum_{n=1}^{n_b}(x_n - S_0 g_n)^2\right) \tag{3.8}$$

with $g_n = (f\exp(-bD^*) + (1-f)\exp(-bD))$.
The unimportant nuisance parameters S_0 and σ_x^2 can be marginalized out. A conjugate normal Gaussian-Inverse-Gamma g-prior is assumed over the nuisance parameters for integration:

$$p(S_0, \sigma_x^2) = \mathcal{N}(S_0|0, \delta^2\sigma_x^2/(\mathbf{g}^T\mathbf{g})) \cdot IG(\sigma_x^2|\alpha, \beta), \tag{3.9}$$

where $\mathbf{g} = [g_1, ..., g_{n_b}]^T$ and IG is a inverse-gamma distribution. The marginalized likelihood is of the form

$$\int_0^\infty \int_{-\infty}^\infty p(S_0, \sigma_x^2) p(\mathbf{x}|f, D, D^*, S_0, \sigma_x^2) dS_0 d\sigma_x^2. \tag{3.10}$$

It can be analytically evaluated by taking the limits of prior parameters $\delta \to 0$ and $\alpha, \beta \to 0$, which encodes a complete lack of prior information. The marginalized likelihood at these limits becomes to

$$p(\mathbf{x}|f, D, D^*) \propto \left[\mathbf{x}^T\mathbf{x} - (\mathbf{x}^T\mathbf{g})/(\mathbf{g}^T\mathbf{g})\right]^{-n_b/2}, \tag{3.11}$$

where proportionality constants are ignored since they do not depend on the parameters of interest.
A hierarchy of prior distribution over the IVIM parameters f, D and D^* is constructed. The parameters are first transformed to incorporate physical plausible parameter domains $0 < f < 1$, $0 < D$ and $0 < D^*$:

$$\begin{aligned} \bar{f} &= \log(f) - \log(1-f) \\ \bar{D} &= \log(d) \\ \bar{D^*} &= \log(D^*). \end{aligned} \tag{3.12}$$

A multivariate Gaussian prior is adopted:

$$p(\boldsymbol{\theta}_i|\boldsymbol{\mu}, \boldsymbol{\Sigma}_{\boldsymbol{\mu}}) = |2\pi\boldsymbol{\Sigma}_\mu|^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\theta}_i - \boldsymbol{\mu})^T\boldsymbol{\Sigma}_\mu^{-1}(\boldsymbol{\theta} - \boldsymbol{\mu})\right) \tag{3.13}$$

with the parameters $\boldsymbol{\theta}_i = \left[\bar{f}_i, \bar{D}_i, \bar{D^*}_i\right]^T$ of voxel i, the ROI mean $\boldsymbol{\mu} = [\mu_f, \mu_D, \mu_{D^*}]^T$ and covariance matrix $\boldsymbol{\Sigma}_\mu$. A noninformative Jeffrey's prior is used for the hyperparameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}_\mu$:

$$p(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\mu) = |\boldsymbol{\Sigma}_\mu|^{-1/2}. \tag{3.14}$$

This prior over hyperparameters has high probability for small covariance determinant, thus outlying estimates are shrunk towards the mean of the distribution.
Voxel-wise IVIM parameters and their summary statistics in a ROI can be estimated simultaneously via the joint posterior, which is given as

$$p(\boldsymbol{\theta}_{1:M}) = \frac{p(\boldsymbol{\mu}, \boldsymbol{\Sigma}_\mu) \prod_{i=1}^{M} p(\mathbf{x}_i|\boldsymbol{\theta}_i) p(\boldsymbol{\theta}_i)}{p(\mathbf{x}_{1:M})} \tag{3.15}$$

with M voxels in the ROI. Estimated IVIM parameters are given by the expectation values under the posterior:

$$\hat{f}_i = \int f_i p(\boldsymbol{\theta}_{i:M}, \boldsymbol{\mu}, \boldsymbol{\Sigma}_\mu|\mathbf{x}_{1:M}) d\boldsymbol{\theta}_{i:M} d\boldsymbol{\Sigma}_\mu d\boldsymbol{\mu} \tag{3.16}$$

and analogously for D and D^*. The integral in Equation (3.16) is analytically intractable and thus a Markov chain Monte carlo (MCMC) method is implemented. IVIM estimators are approximated by the sample statistics from a Markov chain output:

$$\hat{f}_i \approx \frac{1}{N_s} \sum_{j=1}^{N_s} f_i^{(j)} \tag{3.17}$$

with N_s samples and j-th sample $f_i^{(j)}$.

Uncertainty estimation

The inverse mapping of IVIM model is not uniquely solvable due to statistical nature of the model. For instance, accurate estimation with the LSQ approach requires high SNR and precision depends on the parameter combinations [40]. The uncertainty of inverse mapping should provide important information for clinical diagnosis. The uncertainty quantification problem of IVIM inverse mapping has been addressed in [40] with the Cramér-Rao bound (CRB). The Cramér-Rao bound is a useful tool to quantify the uncertainty of parameter estimation [36]. It sets a lower bound on the variance of any unbiased parameter estimators and is given by the inverse of Fisher information matrix $\mathbf{F}$:

$$\begin{aligned} &Cov(\hat{\mathbf{y}}, \hat{\mathbf{y}}) \leq \mathbf{F}^{-1} \\ &\text{with} \quad \mathbf{F} = \mathbb{E}\left[\left(\frac{\partial \ln(\mathbf{L})}{\partial \mathbf{y}} \right)^T \left(\frac{\partial \ln(\mathbf{L})}{\partial \mathbf{y}} \right)^T \right] \end{aligned} \tag{3.18}$$

with the likelihood function $\mathbf{L}$, the parameters to be estimated $\mathbf{y}$ and estimated parameters $\hat{\mathbf{y}}$. However, CRB only gives a theoretical lower bound on the estimation uncertainty of unbiased estimators, which is in practice not achievable. The variance of biased parameter estimators can even be below CRB.
To assess the uncertainty of any IVIM parameter estimator, one can conduct Monte Carlo (MC) simulation similar in [40]. For each parameter combination, numerous noisy and dephased observed signals can be generated according to predefined signal generation procedure and then fitted by an chosen estimator, e.g. the segmented LSQ approach. One can obtain posterior of the estimates, which encapsulates estimation uncertainty of the estimator given the signal model.

3.2 Learning-based Approaches

Either the Cramér-Rao bound or Monte Carlo simulation only provide information about the estimation uncertainty of an estimator, which does not reflect the true uncertainty of inverse mapping. To deal with this problem, we propose to approximate true posteriors using mixture Gaussian model and learn approximated posteriors by neural networks.

3.2.1 Introduction

Signal magnitude is assumed to be normal distributed in the least squares framework and Bayesian shrinkage prior inference described in the last section. However, it is only valid for signals with high SNR and signal averaging breaks the assumption of normality. Likelihood with Rician noise model, signal averaging and dephasing artefacts can be constructed implicitly by defining the sampling procedure for $p(\mathbf{x}|\mathbf{y})$:

$$x_i \leftarrow \frac{1}{n_g} \sum_{j \leq n_g} \text{Rice}(S(b_i; \mathbf{y})\alpha^{\gamma_i}, \nu_j),\ \alpha \sim \mathcal{U}(0,1), \tag{3.19}$$

where n_g is the number of diffusion gradients and $\gamma_i \in \{0,1\}$ is Bernoulli distributed indicator for uniformly distributed attenuation factor α. Each signal component is assumed to be Rician distributed with noise level ν_j and affected by dephasing, which occurs with probability β_i, depending on the b-value [35]. The posterior over model parameters $p(\mathbf{y}|\mathbf{x})$ is defined by the joint distribution over paired data $p(\mathbf{x}, \mathbf{y}) = p(\mathbf{x}|\mathbf{y})p(\mathbf{y})$. We assume uniform distribution for prior $p(\mathbf{y})$ over the physically relevant set of model parameters, similar to the box constrains applied for LSQ in [33].
The posterior $p(\mathbf{y}|\mathbf{x})$ is the object of IVIM parameter estimation. It can be approximated by rejection-based approximate Bayesian computation (ABC) [10]. For each observed $\mathbf{x}$, numerous $\hat{\mathbf{x}}$ are sampled according to Equation 3.19. The generated data is accepted if $D(\mathbf{x}, \hat{\mathbf{x}}) < \epsilon$ with the non-negative acceptance threshold ϵ, where $D(\mathbf{x}, \hat{\mathbf{x}})$ measures the discrepancy between the observed and generated data. For IVIM signals, we define the measure $D(\mathbf{x}, \hat{\mathbf{x}})$ as

$$D(\mathbf{x}, \hat{\mathbf{x}}) = \| \log(\mathbf{x}) - \log(\hat{\mathbf{x}}) \|_1, \tag{3.20}$$

where the log transformation on data is used to eliminate the scale differences of signals with different diffusion encoding strengths. The posterior is approximated by $p(\mathbf{y}|D(\mathbf{x}, \hat{\mathbf{x}}) < \epsilon)$. However, this approximation requires exponential (w.r.t. n_p) number of simulations for each observed signal $\mathbf{x}$, and therefore is infeasible for imaging in practice. Given the efficient sampling procedure defined in Equation 3.19, the posterior can be predicted by neural network optimized stocastically.

3.2.2 Simple Estimation Network

An artificial neural network can be optimized to predict directly IVIM parameters based on sampled paired data $(\hat{\mathbf{x}}, \hat{\mathbf{y}})$ according to Equation 3.19. Such an approach has been proposed in [3] based a simplified signal acquisition model. The estimates are equivalent to the approximation of expectation over the posterior. The network with conditional expectation estimation is referred as Simple Estimation Network (SEN). We adopt a multi-layer perceptron (MLP) with tunabale weights Θ_{SEN}, mapping signal to IVIM parameter space: $\mathbb{R}^{n_b} \rightarrow \mathbb{R}^{n_p}$. The network is optimized by minimizing expected sum of squared error between true and estimated parameter values:

$$\mathcal{L}_{\text{SSE}}(\Theta_{\text{SEN}}) = \mathop{\mathbb{E}}_{p(\mathbf{x},\mathbf{y})} \|\mathbf{y} - \hat{\mathbf{y}}(\mathbf{x}; \Theta_{\text{SEN}})\|_2^2. \tag{3.21}$$

The network architecture is shown in Fig. 3.2.

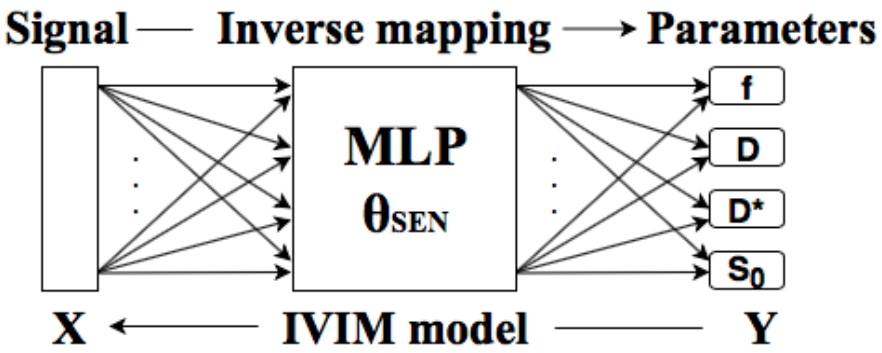


Figure 3.2: Network architecture of SEN.

3.2.3 Amortized Gaussian Posterior

Following the idea of the mixture density network (MDN) [6], full posterior can be modelled as a mixture of Gaussian with mean and diagonal covariance matrices $\{\boldsymbol{\mu}^{(1)}, ..., \boldsymbol{\mu}^{(M)}\}$ and $\{\boldsymbol{\lambda}^{(1)}, ..., \boldsymbol{\lambda}^{(M)}\}$ for M mixture components weighted by the mixture weights $\{\alpha^{(1)}, ..., \alpha^{(M)}\}$:

$$\begin{aligned} p_\Theta(\mathbf{y}|\mathbf{x}) &\triangleq \sum_{m=1}^{M} \alpha^{(m)}(\mathbf{x}; \Theta_{\text{AGP}}) \phi_m(\mathbf{y}|\mathbf{x}; \Theta_{\text{AGP}}) \\ \text{with} \quad \phi_m(\mathbf{y}|\mathbf{x}) &= \mathcal{N}\left(\boldsymbol{\mu}^{(m)}(\mathbf{x}; \Theta_{\text{AGP}}), \text{diag}\left(\boldsymbol{\lambda}^{(m)}(\mathbf{x}; \Theta_{\text{AGP}})\right)\right). \end{aligned} \tag{3.22}$$

These distribution parameters are parametrized by a multi-layer perceptron with tunabale weights Θ_{AGP}, mapping signal to distribution parameter space: $\mathbb{R}^{n_b} \rightarrow \mathbb{R}^{(n_p+2)\times M}$, and reparametrized from network output vector $\mathbf{z}$ by different transformations:

$$\alpha^{(m)} = \frac{\exp(z_m^\alpha)}{\sum_{j=1}^{M} \exp(z_j^\alpha)} \tag{3.23}$$

$$\boldsymbol{\mu}^{(m)} = \mathbf{z}_m^\mu \tag{3.24}$$

$$\boldsymbol{\lambda}^{(m)} = \exp(\mathbf{z}_m^\lambda) \tag{3.25}$$

with z_m^α, $\mathbf{z}_m^\mu$ and $\mathbf{z}_m^\lambda$ corresponding to network outputs used to calculate the weights, mean and covariance of the mixture Gaussian model. Normalization of weights sum to one and non-negativity of covariance are guaranteed in this way.
Minimizing expected KL-divergence between true and variational posteriors $\mathbb{E}_{p(\mathbf{x})} \mathcal{D}_{\text{KL}}(p(\mathbf{y}|\mathbf{x})||p_\Theta(\mathbf{y}|\mathbf{x}))$ is equivalent to maximizing the log-likelihood of predictions:

$$\mathcal{L}(\Theta_{AGP}) = \mathop{\mathbb{E}}_{p(\mathbf{x},\mathbf{y})} \ln \left\{ \sum_{m=1}^{M} \alpha^{(m)}(\mathbf{x}; \Theta_{\text{AGP}}) \phi_m(\mathbf{y}|\mathbf{x}; \Theta_{\text{AGP}}) \right\} - \kappa \sum_{m=1}^{M} \sum_{i=1}^{n_p} \frac{1}{\lambda_i^{(m)}}, \tag{3.26}$$

where the regularization on covariance with the weight κ is used to avoid variance underestimation by maximum likelihood. The network architecture is shown in Fig. 3.3 This network learns full posterior distribution of IVIM parameters, approximated as mixture of Gaussian, hereafter referred as Amortized Gaussian Posterior (AGP).

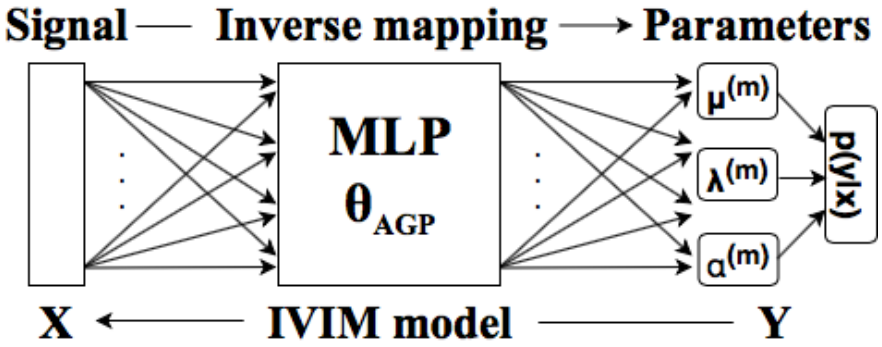


Figure 3.3: Network architecture of AGP.

Chapter 4

Methods: Motion Correction

Motion induced image misalignment is hard to be included into the IVIM inverse mapping procedure, especially when the mapping is voxel-wise. None of the fitting algorithms described in the last section is invariant to that. Therefore, image misalignment needs to be corrected in practice before inverse mapping to improve estimation accuracy. The following section describes the proposed motion correction pipeline including image averaging and 3D volume registration.

The fetal MR dataset consists of 3D volumes with different gradient encoding directions and strengths. We align the images before model fitting in two steps: 1) construct high SNR volumes by averaging the images with the same b-value, 2) register the averaged 3D volumes with different b-values. For both steps, we apply a group registration algorithm to address the issues of contrast changes and dephasing artefacts. The images with the same b-value are aligned slice-by-slice, whereas the images with different b-values are aligned volume-wise.

In the group registration framework, all the images $\mathbf{I} = \{\mathbf{I}_1, ..., \mathbf{I}_G\} : \Omega \to \mathbb{R}$ with a discrete image domain Ω are registered simultaneously to a virtual reference space. It can be formulated as the following optimization problem, which estimates G N-dimensional displacement fields $\mathbf{d} = \{\mathbf{d}_1, ..., \mathbf{d}_G\}$ by minimizing the cost function $\ell(\mathbf{d})$:

$$\mathbf{d}^* = \underset{\mathbf{d}}{\operatorname{argmin}}\, \ell(\mathbf{d}) \tag{4.1}$$

$$= \underset{\mathbf{d}}{\operatorname{argmin}}\, E_{\mathrm{D}}(\mathbf{d}; \mathbf{I}) + \lambda E_{\mathrm{R}}(\mathbf{d}) \tag{4.2}$$

with an image dissimilarity metric E_{D}, displacement regularization term E_{R} and regularization weight λ. The images containing L pixels and corresponding displacement fields are considered as vectors: $\mathbf{I}_g \in \mathbb{R}^L$ and $\mathbf{d}_g = (\mathbf{d}_{g,1}^{\mathsf{T}}, ..., \mathbf{d}_{g,N}^{\mathsf{T}}) \in \mathbb{R}^{LN}$ with the n-th component $\mathbf{d}_{g,n}$ of the displacement field $\mathbf{d}_g$.

We use a principle component analysis (PCA) based groupwise dissimilarity metric proposed in [14]. We define an $L \times G$ matrix $\mathbf{M}$ which contains the transformed images as columns:

$$\mathbf{M} = [\, g(\mathbf{I}_1, \mathbf{d}_1) \quad ... \quad g(\mathbf{I}_G, \mathbf{d}_G) \,] \tag{4.3}$$

with the warping function $g(\cdot, \cdot)$, which warps the image $\mathbf{I}_g$ with the displacement field $\mathbf{d}_g$. The dissimilarity metric is defined as the nuclear norm of the image matrix $\mathbf{M}$, which is equal to the sum of the singular values σ of $\mathbf{M}$:

$$E_D = \| \mathbf{M} \|_* = \sum_{i=1}^{G} \sigma_i(\mathbf{M}). \tag{4.4}$$

This metric is based on the assumption that aligned images are linearly correlated, while misaligned images contain more complex data. Since the size of fetal organs is rather small compared to the whole MR image, contrast changes can happen inside a tiny area. It is hard to capture these changes while considering the whole image as an unit. Therefore, we divide each image into non-overlapping patches with equal size and apply the nuclear metric on each patch. The dissimilarity metric becomes to be the sum of the nuclear metrics in each patch:

$$E_D = \sum_{p=1}^{P} \| \mathbf{M}(\mathbf{I}^p, \mathbf{d}^p) \|_*, \tag{4.5}$$

where P is number of patches, $\mathbf{I}^p$ and $\mathbf{d}^p$ represents the images and displacement fields in the p-th patch respectively.
For the regularization term, we employ a sparse and non-convex spatial ℓ^p regularization with $0 < p < 1$ based on displacement gradients to enforce smoothness of displacement fields:

$$E_R(\mathbf{d}) = v\|D(\mathbf{d})\|^p \tag{4.6}$$

$$D(\mathbf{d}) = [\nabla_1 d_1 \ \nabla_2 d_1 \ \ldots \ \nabla_1 d_2 \ \ldots \ \nabla_N d_N]^T \in \mathbb{R}^{N^2 \times L}, \tag{4.7}$$

where v is the physical pixel volume and $\nabla_i \in \mathbb{R}^{L \times L}$ is the derivative operator along the i-th component.
Following [38], the displacement fields $\mathbf{d}$ are parametrized using cubic interpolation with the displacements $\mathbf{k}$ on regularly spaced control points to reduce the dimensionality of the registration optimization problem. It is sufficient to impose regularization on the displacements $\mathbf{k}$. The image registration problem with parametric displacement fields is formulated as

$$\mathbf{k}^* = \underset{\mathbf{k}}{\operatorname{argmin}}\, \ell(\mathbf{d}(\mathbf{k})) \tag{4.8}$$

$$= \underset{\mathbf{k}}{\operatorname{argmin}}\, E_{\mathrm{D}}(\mathbf{d}(\mathbf{k}); \mathbf{I}) + \lambda\eta\|D(\mathbf{k})\|^p \tag{4.9}$$

with the physical volume of the control grid η. This registration optimization problem is solved by an efficient numerical algorithm using the Alternating Directions Method of Multiplier (ADMM) proposed in [38].

Chapter 5

Experiments and Results: Simulation

5.1 Experiment Setup

5.1.1 Simulation Data

Isotropic data

Signals are simulated according to Equation (3.19). We assume uniform prior distribution $p(\mathbf{y})$ with the following parameter ranges: $S_0 \in [0, 3000]$, $f \in [0.0005, 0.9995]$, $D \in [0.045, 5]\times10^{-3}\text{mm}^2/\text{s}$ and $D^* \in [0.34, 100]\times10^{-3}\text{mm}^2/\text{s}$. We dismiss samples with $D>D^*$. In accordance with in vivo measured fetal MR images described in the next chapter, we use Rician noise with $\nu \in [6, 9]$ for the diffusion encoding gradient $[0, 0, 0]$ and $[15, 18]$ for the rest. The noise level is estimated by extracting a background region from images and fit a Rician distribution onto its histogram. The different noise levels are caused by the use of tetrahedral gradients [9] and the image reconstruction procedure of the scanner. The dephasing probability γ_i is chosen according to the in vivo data quality and is b-value dependent: 2% for $b<300\,\text{s}/\text{mm}^2$ and linearly increased from 10% to 25% for $b\geq300\,\text{s}/\text{mm}^2$. To construct trace signals, we average the signals with four different noise, which is encoding direction dependent, and dephasing realizations.

Anisotropic data

Original IVIM model [21] considered in Equation (3.1) is based on the fact that rotation invariant trace signal of measurements with orthogonal gradient directions approximates directionally averaged contributions of diffusion and perfusion. The anisotropic property of diffusion and perfusion can be more accurately reflected by tensor modeling [12]. We simulate 1024 randomly rotated diffusion tensors with $\text{tr}(\mathbf{D})=9.4\cdot10^{-4}\text{mm}^2/\text{s}$, $\text{tr}(\mathbf{D}^*)=5.3\cdot10^{-2}\text{mm}^2/\text{s}$, fractional anisotropy of 0.8 and $f=0.18$. The observed signal is simulated using tetrahedral diffusion gradient [9] applied in the in vivo measurement. Rician noise, dephasing and signal averaging are applied to the simulated signal according to Equation (3.19). We take the trace of perfusion and diffusion tensor as ground truth ($D\approx\text{tr}(\mathbf{D})$).

5.1.2 Hardware and Software Framework

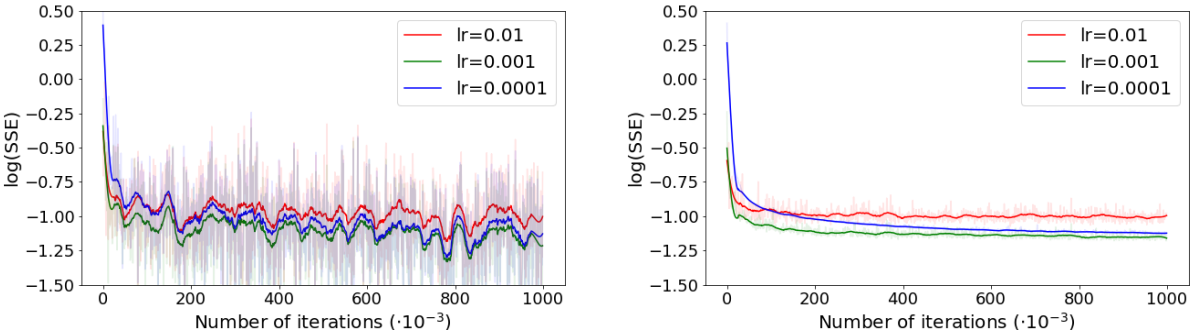
We use the registration toolbox from [38] with the code available at https://github.com/visva89/pTVreg. The Bayesian shrinkage prior inference used the Matlab script from [33]. Artificial neural networks (ANN) are implemented using Tensorflow [1]. All ANN-related experiments are performed on GEForce GTX TITAN X GPUs.

5.2 Simple Estimation Network

We adopt a multilayer perceptron to estimate IVIM parameters. In the following sections, we explore the influence of network capacity and choice of hyperparameters on estimation accuracy for different simulation settings. Simulated trace signals from isotropic data are used for training and validation. Training data are generated on-the-fly during training, whereas a fix amount (10^4) of signals is used for validation. To reduce bias caused by different parameter scaling, network input (signals) and output (IVIM parameters) are standardized. The standardization parameters mean and variance are calculated from 10^7 randomly generated samples according to Equation (3.19).
In all experiments, hyperbolic tangent activation is used at each layer except the output layer, which is linear. The weights in MLP are initialized as samples from a truncated Gaussian distribution of zero mean and variance $\frac{2}{N_{in}+N_{out}}$ with number of input neurons N_{in} and output neurons N_{out}. The network is trained for 10^6 iterations using Adam optimizer [17]. We determine the optimal structure for MLP in Fig. 3.2 based on the learning curves of training and validation data.

5.2.1 Learning rate

A proper learning rate for optimizing the IVIM model fitting network needs to be determined before further exploring other hyperparamters. We search across the learning rates coarsely. A multilayer perceptron with three hidden layers and 20 nodes per layer is trained with a batch size of 100 samples and learning rates $lr = 0.01$, $lr = 0.001$ and $lr = 0.0001$ for simulated signals without dephasing. The learning curves for training and validation data are depicted in Fig. 5.1. We report the mean loss (sum of squared error between true and estimated parameters) in each batch of training data and in the whole validation dataset at every 1000 iterations. The curves, displayed in logarithmic scale, show that the training of this network is not sensitive to learning rate. In all three cases, training converges after roughly half of the total iterations. The network trained with $lr = 0.01$ has slightly higher loss, whereas the networks with $lr = 0.001$ and $lr = 0.0001$ yield comparable results. In addition, no overfitting is observed here, since the training and validation data are sampled from the same distribution. We fix learning rate to 0.001, which allows faster convergence, for the rest of experiments.



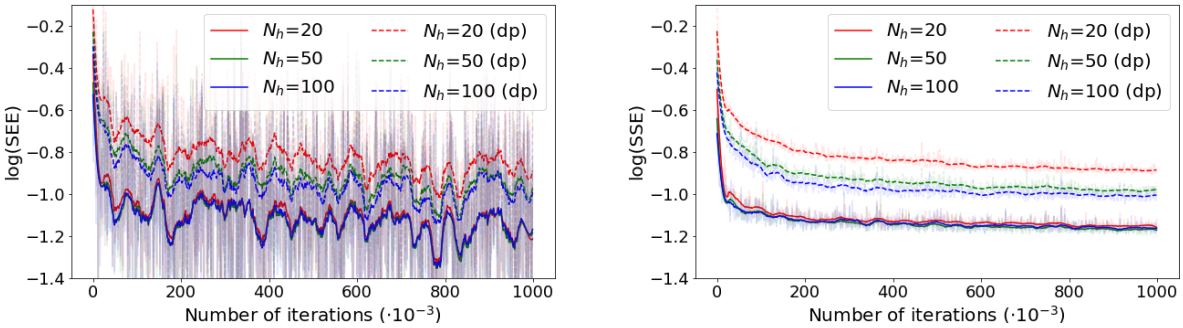
(a) Training data

(b) Validation data

Figure 5.1: Learning curves with different learning rates (lr) for (a) training data and (b) validation data. The loss, sum of squared error (SSE), is displayed in the logarithmic scale. Solid lines represent the learning curves smoothed by the Savitzky-Golay filter. Transparent lines are the original learning curves.

5.2.2 Network capacity

Starting from the basic MLP used in the last section, we experiment with different network width and depth to determine required capacity for this inverse mapping problem. For those experiments, the batch size is set equal to 100. We first consider the number of nodes in hidden layers for simulation without and with dephasing. Again the training and validation samples are from the same distribution. Learning curves for three different number of nodes $N_h = 20, 50, 100$ are shown in Fig. 5.2. Two observations stand out from the figures. First, the network, estimating parameters from dephased signals, has lower estimation accuracy. Second, there is a limit amount of performance we can achieve by increasing the number of nodes. The performance saturates when $N_h = 20$ for simulation without dephasing and $N_h = 50$ for simulation with dephasing. Therefore, we set the number of nodes equal to 50.



(a) Training data

(b) Validation data

Figure 5.2: Learning curves with different number of nodes in the hidden layers N_s for (a) training data and (b) validation data. The loss, sum of squared error (SSE), is displayed in the logarithmic scale. Solid lines represent the learning curves smoothed by the Savitzky-Golay filter. Transparent lines are the original learning curves.

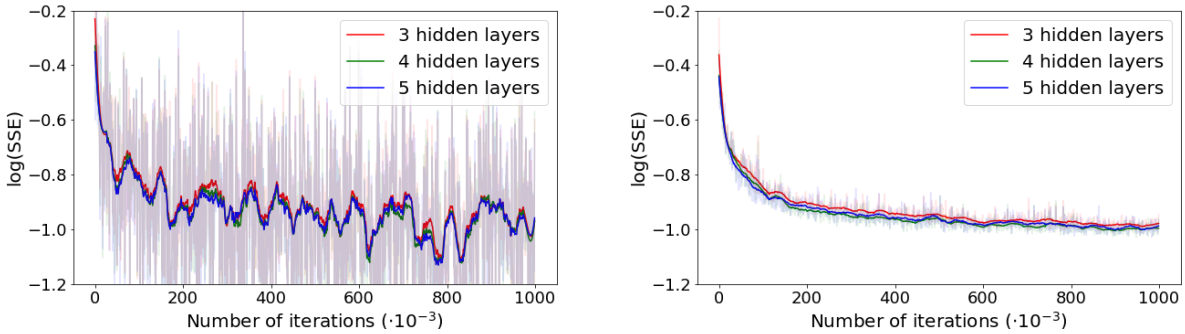
Next we explore different network depth for simulation with dephasing. Fig. 5.3 depicts the learning curves for the networks with 3, 4 and 5 hidden layers and 50 nodes per layer. It can be observed that increasing the network depth does not improve the performance substantially.

5.2.3 Batch size

The batch size is determined for the network with 5 hidden layers and 50 nodes per layer trained with simulated data with dephasing. From the learning curves with $B_s = 100, 1000, 2000$ depicted in Fig. 5.4, we can observe that larger batch size does not only improve the performance, but also reduce the error fluctuation during training. Smaller batch can lead to less accurate estimate of the gradients. Still there is a limit of performance improvement brought by large batch size. We set the batch size to 2000 for the rest of experiments.

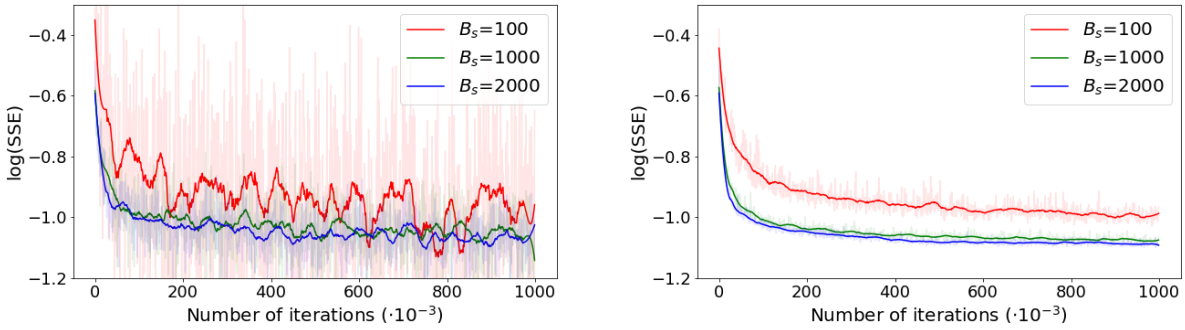
5.3 Amortized Gaussian Posterior

We adopt the training settings of SEN determined in the last section for AGP. We show here that AGP requires more network capacity to store information for inference, since the optimization problem becomes



(a) Training data (b) Validation data

Figure 5.3: Learning curves with different number of hidden layers for (a) training data and (b) validation data. The loss, sum of squared error (SSE), is displayed in the logarithmic scale. Solid lines represent the learning curves smoothed by the Savitzky-Golay filter. Transparent lines are the original learning curves.



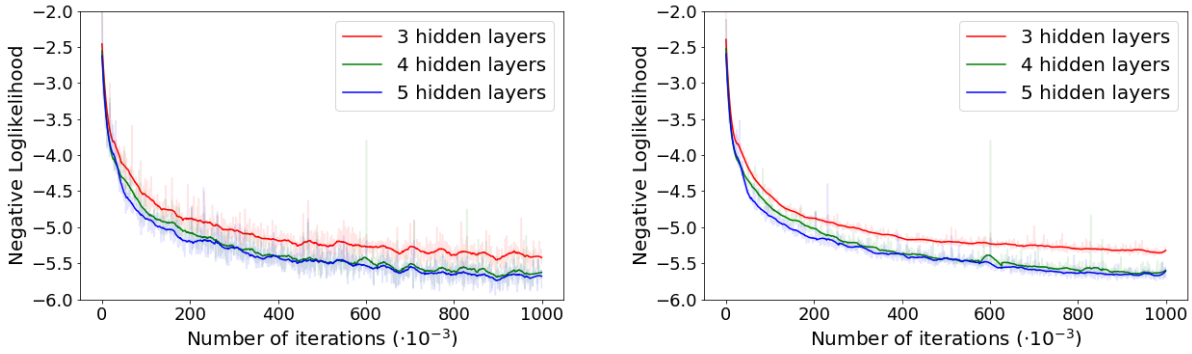
(a) Training data (b) Validation data

Figure 5.4: Learning curves with different batch sizes B_s for (a) training data and (b) validation data. The loss, sum of squared error (SSE), is displayed in the logarithmic scale. Solid lines represent the learning curves smoothed by the Savitzky-Golay filter. Transparent lines are the original learning curves.

more difficult to solve. A mixture density network with one mixture component, varied number of hidden layers and 50 nodes per layer is trained with a batch size of 2000 samples. Fig. 5.5 shows the learning curves of training and validation data. We can observe that the use of more hidden layers is evidently beneficial. In addition, the network converges much slower than SEN shown in Fig. 5.3

5.4 Simulation Results

To carry out a fair comparison among the learning-based approaches, all the networks have the same capacity (5 hidden layers and 50 nodes per layer), trained for 10^6 iterations with a learning rate of 0.001 and a batch size of 2000 samples. We compare the estimation performance on the simulated isotropic data of LSQ, SEN and three variants of AGP: AGP with one mixture component AGP-1, two mixture components AGP-2 and AGP-1 with regularized variance AGP-1-reg. Estimated IVIM parameter value by AGP-1 and AGP-1-reg is



(a) Training data

(b) Validation data

Figure 5.5: Learning curves with different number of hidden layers for (a) training data and (b) validation data. Solid lines represent the learning curves smoothed by the Savitzky-Golay filter. Transparent lines are the original learning curves.

equal to the predictive mean of the Gaussian posterior. For AGP-2, we calculate the mode of the predictive mixture Gaussian distribution as the estimated parameter value.

5.4.1 Simulated Isotropic Data

Estimation accuracy

We generate 10^7 samples according to Equation 3.19 with dephasing for test purpose. Mean and median absolute errors between the estimated and true parameters are reported in Table 5.1. In results, the learning based approaches SEN and AGP outperform LSQ substantially in estimating all four parameters, whereas the three AGP variants achieve lower estimation error than SEN. In addition, we can observe that the difference between the mean and median values of absolute errors is large, especially in f, D and D^*, indicating the existence of samples with high estimation difficulty.

Table 5.1: Mean and median absolute errors between estimated and true IVIM parameter values by LSQ, ANN, AGP-1, AGP-1-reg and AGP-2 on the simulated isotropic data. Bold number indicates the lowest error.

	S_0		f [%]		D [10^{-4} mm/s^2]		D^* [10^{-3} mm/s^2]	
methods	mean	median	mean	median	mean	median	mean	median
LSQ	98.78	45.79	9.25	4.45	6.80	3.11	22.68	12.59
SEN	14.34	11.16	3.79	1.26	2.25	0.86	5.64	1.93
AGP-1	11.33	**8.57**	3.49	0.78	**2.21**	**0.75**	5.68	1.79
AGP-1-reg	11.48	9.00	3.50	0.79	2.22	0.79	**5.61**	1.77
AGP-2	**11.09**	**8.57**	**3.43**	**0.75**	2.26	0.76	5.98	**1.66**

Uncertainty quantification

In the following experiment, we compare posteriors of the estimates by the segmented LSQ method via Monte Carlo simulation with sampled posteriors on the isotropic data. For a given parameter set, signals with 10^5 different noise and dephasing realizations are generated according to Equation 3.19. LSQ posteriors are constructed from the fitted parameters by LSQ to the generated signals. Two sample scenarios are depicted in Fig. 5.6, showing the comparison between the sampled posteriors and the results of MC simulation for LSQ. The top row in Fig. 5.6 depicts a parameter combination where perfusion and diffusion effects can be separated with less uncertainty. The bottom row in Fig. 5.6 indicates high uncertainty in estimating D^*. This can be caused by a low perfusion fraction f, which makes the perfusion and diffusion effects indistinguishable from each other. We can observe that the posteriors of LSQ have larger variance than the sampled posteriors. There are peaks on the parameter range boundaries in the posterior plots for D^* and f, indicating possible values outside the pre-defined parameter ranges.

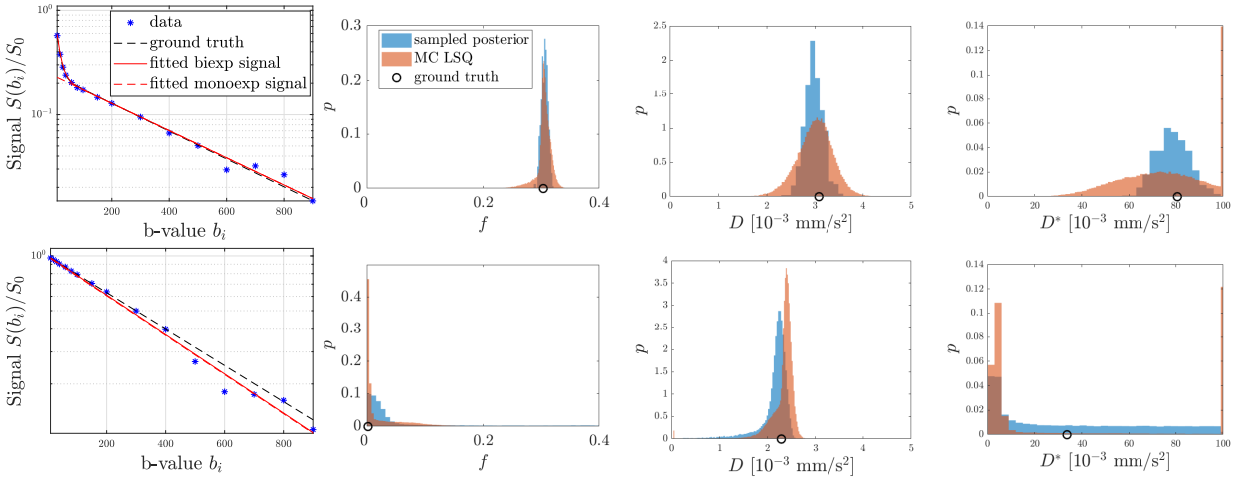


Figure 5.6: Sampled posteriors and Monte Carlo simulation results of LSQ for two simulated signals.

Next we verify the inverse mapping inferred by AGP on the isotropic data. We report the estimation uncertainty as the corresponding standard deviation of predictive posterior. The same sample scenarios as in Fig. 5.6 are depicted in Fig. 5.7, showing the comparison among the estimated posteriors by three AGP variants. We can observe underestimation of the predictive variance by AGP-1. This can be partially alleviated by applying regularization, evident in the posterior for f in the top row. However, AGP-1 and AGP-1-reg yield very similar results in most cases shown here. Smaller estimated variance than true variance, could also arise from the approximation error caused by the choice of rejection threshold ϵ. A large threshold can include more data samples and lead to wider posterior distribution. Although sampled posteriors can be well approximated by simple Gaussian distribution, the posteriors inferred by AGP-2 are in better agreement with the sampled posterior for D^* in both scenarios and for f in the bottom row, indicating potential estimation improvement by using more mixture components.

Inverse mapping uncertainty of IVIM parameters f, D and D^* with varying simulated parameter combinations are depicted in Fig. 5.8. In each experiment, one parameter is fixed while the other two are varied in 1000 steps within the ranges defined for the signal acquisition model. Each estimation uncertainty is the averaged standard deviation of predictive posteriors over 1000 different noise realizations and without dephasing. S_0 is set to 1000. In results, high uncertainty is observed when D and D^* values close to each other or when f approaches one or zero. For instance, the estimation of f has high uncertainty for f close

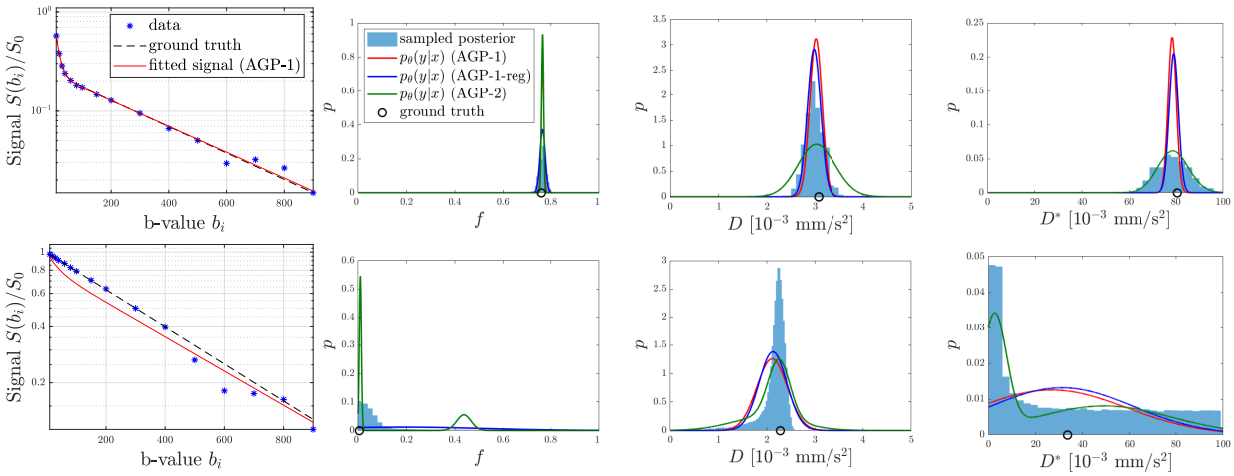


Figure 5.7: Sampled posterior and predictive posteriors by AGP-1, AGP-1-reg, AGP-2 for two simulated IVIM signals.

to zero and similar D and D^* values. The behavior of the estimation uncertainty for D^* is well illustrated along the red dotted line in the bottom-right plot in Fig. 5.8: The uncertainty increases with larger values of D^*, since the large perfusion decreases the signal $\exp(-bD^*)$ to zero; while, a high uncertainty is observed for smaller values of D^* as well, but this time due to similar perfusion and diffusion contribution.

5.4.2 Simulated Anisotropic Data

We compare SEN, AGP-1 with the segmented LSQ approach on the simulated anisotropic data to test the sensitivity of the fitting methods to modelling assumption. Mean absolute errors between the estimated and true parameters are reported in Table 5.2, showing that AGP-1 achieves the smallest estimation errors in all three parameters for the signals modeled both with and without dephasing. For signals with dephasing, SEN and AGP-1 outperforms the LSQ approach substantially.

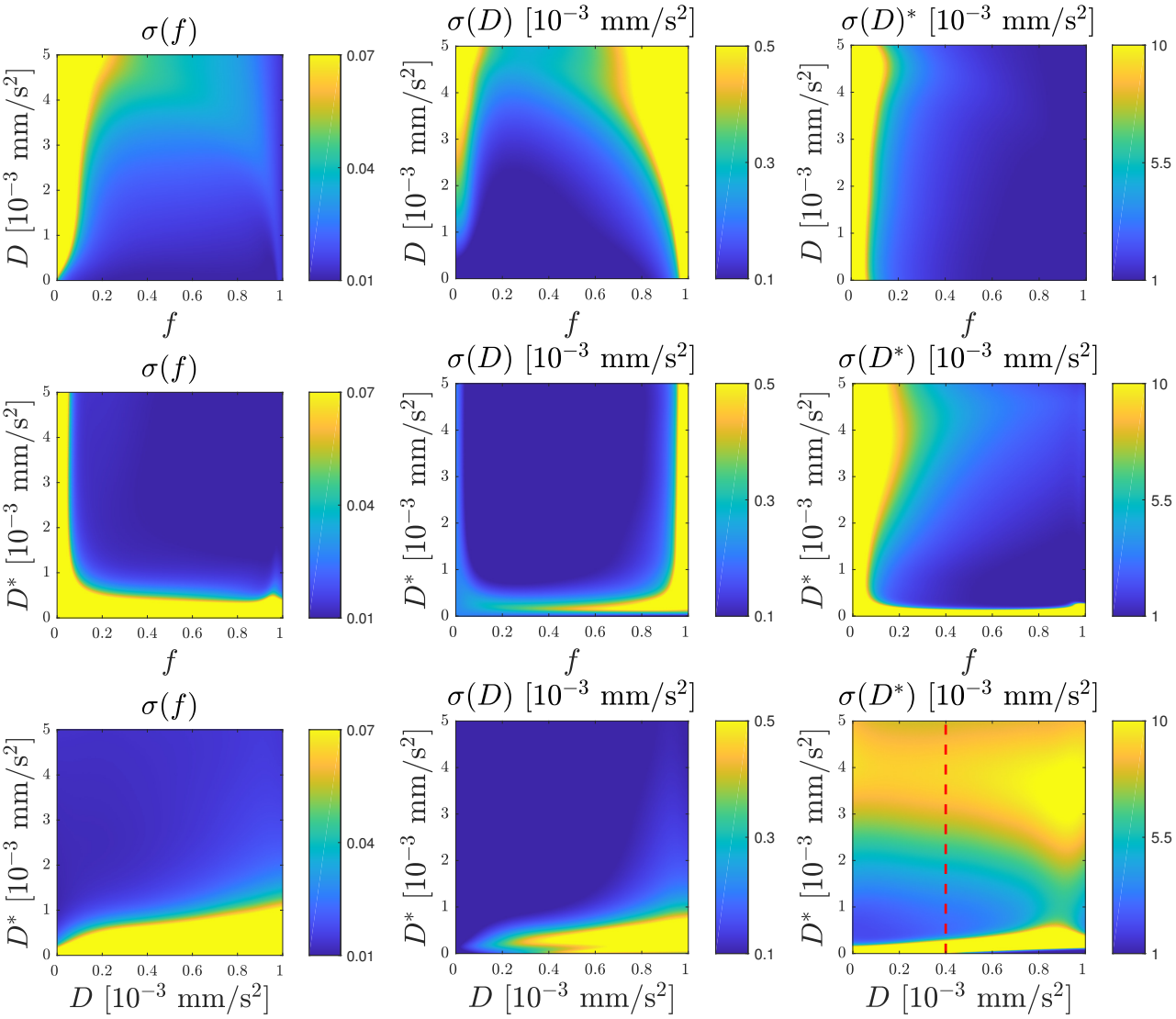


Figure 5.8: Estimation uncertainty with different parameter combinations based on simulation without dephasing inferred by AGP-1. σ is the standard deviation of the predictive posterior. Top row: D^*=0.02 mm^2/s, middle: D=0.001 mm^2/s, bottom: f=0.2. S_0=1000.

Table 5.2: Estimated mean IVIM parameter values with LSQ, SEN and AGP-1 on the simulated anisotropic data with and without dephasing (dp). Mean absolute error (MAE) is calculated between the ground truth and estimated parameters. Bold number indicates the lowest error.

			LSQ	**SEN**	**AGP-1**
Mean	f [%]	w/o dp	17.9	19.2	19.0
		w. dp	17.6	18.9	18.8
	D [10^{-4} mm/s^2]	w/o dp	7.74	8.24	8.72
		w. dp	9.09	9.13	9.16
	D^* [10^{-3} mm/s^2]	w/o dp	29.3	40.7	43.1
		w. dp	44.5	44.4	46.2
MAE	f [%]	w/o dp	2.6	2.6	**2.2**
		w. dp	8.4	2.9	**2.5**
	D [10^{-4} mm/s^2]	w/o dp	1.69	2.13	**1.64**
		w. dp	2.27	1.38	**1.30**
	D^* [10^{-3} mm/s^2]	w/o dp	29.3	13.3	**11.2**
		w. dp	40.1	11.1	**9.7**

Chapter 6

Experiments and Results: In-vivo Fetal MR

6.1 Experiment Setup

6.1.1 In-vivo Data

Prenatal MRI was acquired in 17 subjects on a 1.5 T Discovery MR450 unit (GE Healthcare, Milwaukee, WI, USA). A dual spin-echo planar sequence was performed with echo time of $2200/75$ ms, acquisition matrix 80×100, voxel size of 2×2 mm^2, slice thickness 3 or 4 mm and slice number of 9-24. The tetrahedral diffusion gradients [9] were used with 16 b-values: 10, 20, 30, 40, 60, 80, 100, 150, 200, 300, 400, 500, 600, 700, 800, 900 mm/s^2, yielding 64 diffusion weighted images for one subject. One b_0 image was acquired. The acquisition plane is axial for fetal brain imaging and coronal for other organs and placenta. The IVIM imaging sequence was repeated twice for each subject.

6.1.2 Evaluation Criteria

For evaluating the estimation performance on the in-vivo data, we use two metrics (1) root mean squared error (RMSE) between the fitted and observed data points to quantify goodness of fit, (2) intra-subject variation VAR% to measure rescan repeatability, defined as

$$\text{VAR}\% = \frac{100}{N} \sum_{i=1,\ldots,N} \frac{|\text{TEST}_i - \text{RETEST}_i|}{|\text{TEST}_i + \text{RETEST}_i|/2} \tag{6.1}$$

with the number of subjects N, repeated acquisitions TEST_i and RETEST_i of i-th subject. TEST_i and RETEST_i represents the mean of the parameters in the manually segmented region of interest.

6.2 In-vivo Results

The images with different b-values from two subjects are shown in Fig. 6.1. The top row shows an example with well imaged placenta, brain and fluid tissues, whereas the bottom row illustrates an example with focus on brain, lung and kidney. We can observe from the figures that fetal organs show different image contrast while diffusion encoding strength changes. The images are overlaid with manually segmented mask for placenta by an expert and the markers on the images serve for the illustration of the fitting results coming after.

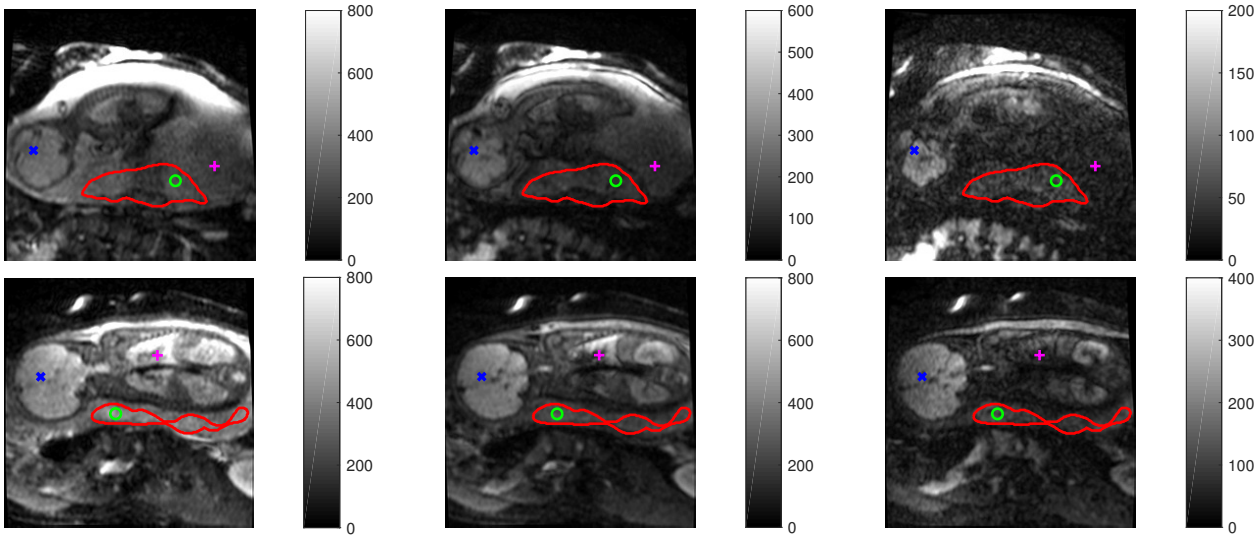
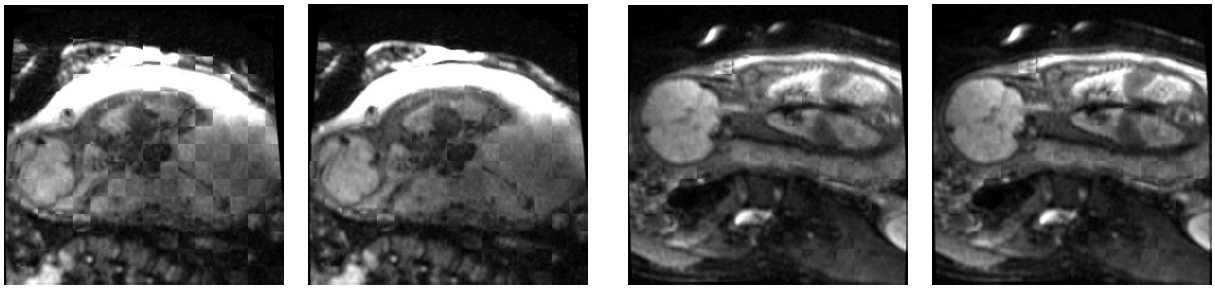


Figure 6.1: Two in-vivo fetal MR examples. Aligned images with three different b-values are shown: $b = 0$ (left), $b = 200$ (middle) and $b = 700$ (right).

Registration results

We use a parameter configuration with norm $p = 0.8$, control point grid spacing $K = 10$, regularization weight $\lambda = 0.01$ for 2D registration and $\lambda = 0.008$ for 3D registration. Examples of the registration results are illustrated in Fig. 6.2, showing the checkerboard compositions of two images ($b = 10$ and $b = 40$) with alternating blocks prior to and after registration for the in vivo examples shown in Fig. 6.1. The comparison between the image compositions in (a) with large motion shows that the images are aligned better after registration.



(a) (b)

Figure 6.2: Registration results (a) corresponds to the example in the top row, (b) in the bottom row, in Fig. 6.1. In each subplot, the left image is prior to registration, whereas the right image is after registration.

Importance of dephasing simulation

To assess the importance of dephasing simulation for the learning-based fitting algorithms, we compare SEN with its variant SEN-n trained on the simulated data without dephasing. Fig. 6.3 illustrates the fitted signals

for the marked voxels in Fig. 6.1, showing that dephasing artefacts have large impact on SEN-n and the fitted curves are biased by the dephased data points.

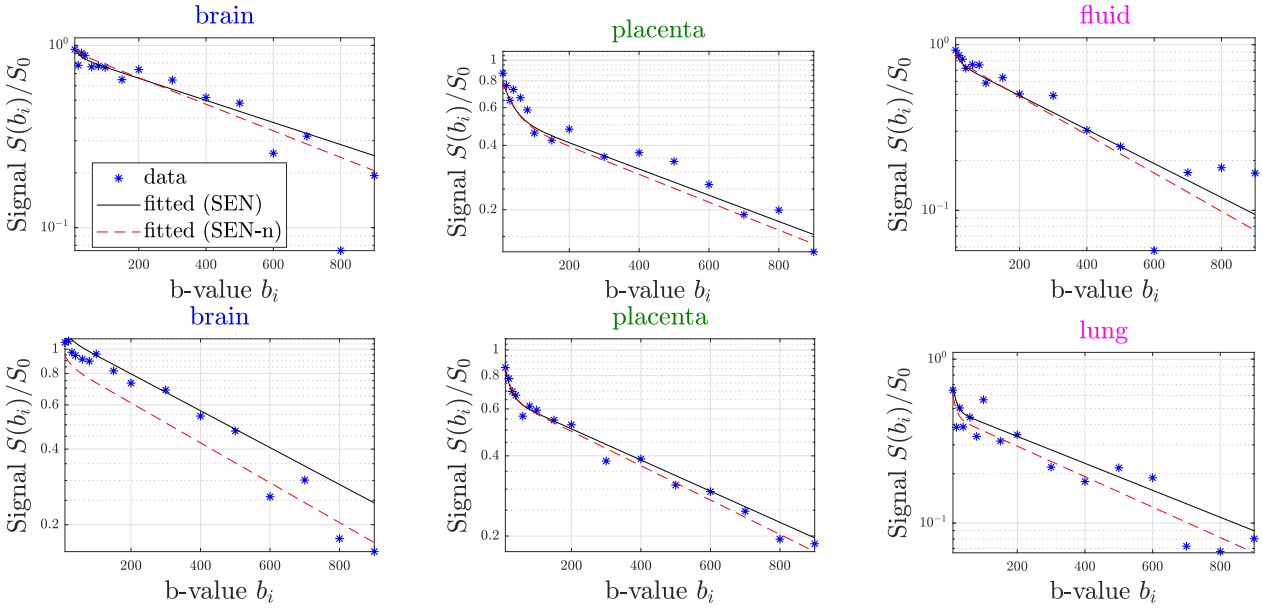


Figure 6.3: Fitted signals by SEN and SEN-n for the voxels marked in Fig.6.1. The top row corresponds to the example in the top row of Fig.6.1, analogous for the bottom row.

Fitted parameter maps

Estimated parameter and RMSE maps by LSQ, BSP, SEN and AGP-1 are shown in Fig. 6.5 and Fig. 6.6 for the two subjects in Fig. 6.1, respectively. Only estimates in the delineated placenta region are shown for BSP. The maps given by SEN and AGP-1 have more homogeneous appearance, while the fitted maps in the placenta by BSP look patchy. Since placenta is a multi-compartment organ containing highly perfused maternal and fetal vasculature, the approximation of prior distribution as a multivariant Gaussian in BSP does not hold anymore. Inappropriate choice of prior distribution has huge impact on the estimation performance of the Bayesian approach. The poor estimation accuracy of BSP on fetal placenta is also reflected by substantially high RMSE compared to the another three methods in both in-vivo examples. However, adapting prior modelling to the anatomical structure of placenta would improve esitmation performance in a significant way. LSQ is less reliable in estimating D^*, since there are a large amount of outlier pixels, which have values on the parameter boundaries. In addition, the D^* values estimated by LSQ are much lower than SEN and AGP-1.

The uncertainty maps provided by AGP-1 reveal that inverse mapping is less precise in the regions with severe dephasing artefacts, e.g. brain with a lot signal loss due to fetal head motion. Signals with low SNR, which are typically in the highly perfused regions, cannot be estimated reliably as well. The three fitted signals in Fig. 6.5 and Fig. 6.6 show that the dephased data points at high b-values bias the fitting of LSQ, but SEN and AGP-1 are less affected by them. Better explanation of perfusion effect in brain, placenta and fluid by SEN and AGP-1 are evident in the fitted signals. Small anatomical structures, e.g. lung and rib, pose large difficulty for registration due to their size and contrast changes. Image misalignment affects the

fitting performance of SEN and AGP, since the proposed implicit signal acquisition model used for training cannot explain misaligned data. The fitted curves for a voxel in lung shown in Fig. 6.6 further illustrate the unreliability of LSQ, SEN, AGP-1 in parameter estimation for misaligned data. Moreover, comparing to the estimation uncertainty illustrated in Fig. 5.8, the parameter estimation for the in vivo data has much higher uncertainty, indicating that reliable estimation is hard achievable due to poor data quality.
According to the RMSE maps, LSQ achieves the lowest residuum between fitted and observed signals among all the fitting approaches. This can be explained by the fact that LSQ is optimized for minimizing the L_2 norm. SEN and AGP-1 take dephasing and Rician noise into account, which can lead to higher RMSE. This is especially evident in the brain with severe dephasing artefacts. However, AGP-1 achieves lower RMSE compared to SEN in both in-vivo examples, pointing out the existence of better fitting even while considering dephasing artefacts. It is in agreement with the simulation results, where AGP outperforms SEN with respect to estimation accuracy.

Repeatability experiments

Repeatability results for the placenta, the mean and standard deviation of the average parameter values in the ROI across the subjects and the RMSE between fitted and observed signals are reported in Table 6.1 for LSQ, BSP, SEN and AGP-1. Registration improves repeatability in D and D^*, but increases variability in f slightly with all four fitting methods. Lower RMSE is achieved by registration, but can be explained by image smoothing caused by cubic interpolation in the registration procedure. Overall, there is no obvious improvement with respect to paired rescan repeatability brought by motion correction.
However, substantial decrease of intra-subject variation in all three parameters can be achieved by both learning-based methods SEN and AGP-1. It indicates that intra-subject variability could mainly arise from the limitation of LSQ fitting approach in low SNR scenarios and its instability to dephasing artefacts.
The estimated mean values for D^* by SEN and AGP-1 is on average 25% higher than LSQ, indicating estimation with more perfusion effect by the learning-based approaches. It is in accordance with the fitted signals in Fig. 6.5 and Fig. 6.6, showing perfusion at low b-values is better explained by AGP-1 and SEN. Moreover, the motion correction procedure decreases the mean values for D^* in all estimation methods. One reason could be that image misalignment at low b-values can cause false interpretation of signal fluctuation as perfusion effect.

Table 6.1: Fetal in vivo intra-subject repeatability of IVIM mapping inside placenta. The smallest variance and RMSE indicated in bold.

Method		**VAR%**			**Mean**			**Standard Deviation**			**RMSE**
		f	D	D^*	f[%]	D[mm^2/s]	D^*[mm^2/s]	f[%]	D[mm^2/s]	D^*[mm^2/s]	
LSQ	w/o reg	23.17	12.47	36.64	30.5	$1.51{\cdot}10^{-3}$	$3.03{\cdot}10^{-2}$	8.3	$0.26{\cdot}10^{-3}$	$1.01{\cdot}10^{-2}$	30.72
	w. reg	24.92	11.37	34.35	30.7	$1.51{\cdot}10^{-3}$	$2.63{\cdot}10^{-2}$	7.6	$0.23{\cdot}10^{-3}$	$0.90{\cdot}10^{-2}$	**26.34**
BSP	w/o reg	20.54	61.79	64.75	44.2	$5.90{\cdot}10^{-3}$	$1.44{\cdot}10^{-2}$	14.1	$6.34{\cdot}10^{-3}$	$1.66{\cdot}10^{-2}$	61.98
	w. reg	20.87	48.95	50.20	39.9	$4.20{\cdot}10^{-3}$	$0.85{\cdot}10^{-2}$	12.6	$2.77{\cdot}10^{-3}$	$0.63{\cdot}10^{-2}$	59.26
SEN	w/o reg	**11.03**	12.17	13.87	34.0	$1.39{\cdot}10^{-3}$	$3.74{\cdot}10^{-2}$	6.0	$0.24{\cdot}10^{-3}$	$0.80{\cdot}10^{-2}$	42.48
	w. reg	11.26	10.61	**11.75**	33.8	$1.40{\cdot}10^{-3}$	$3.57{\cdot}10^{-2}$	5.1	$0.21{\cdot}10^{-3}$	$0.77{\cdot}10^{-2}$	38.40
AGP-1	w/o reg	11.58	12.17	15.60	32.9	$1.43{\cdot}10^{-3}$	$3.82{\cdot}10^{-2}$	6.5	$0.25{\cdot}10^{-3}$	$0.86{\cdot}10^{-2}$	36.55
	w. reg	12.24	**10.60**	13.63	32.4	$1.46{\cdot}10^{-3}$	$3.68{\cdot}10^{-2}$	5.7	$0.21{\cdot}10^{-3}$	$0.81{\cdot}10^{-2}$	32.24

High bias of the fitting methods can lead to low intra-subject variation as well. Next, we examine the bias of

LSQ, SEN and AGP-1 by exploring the histograms of the fitted parameters across subjects. The histograms are depicted in Fig. 6.4, showing the distribution of fitted parameters in the central slice of all the subjects. Background voxels and outlier voxels of LSQ, which have values on the parameter box constraints, are not included in the histograms. No large bias in any IVIM parameters is observed for SEN and AGP-1, whereas the D^* histogram of LSQ shrinks towards low apparent diffusion coefficient. It reveals the underestimation of perfusion effect by LSQ, possibly due to dephasing artefacts at high b-values.

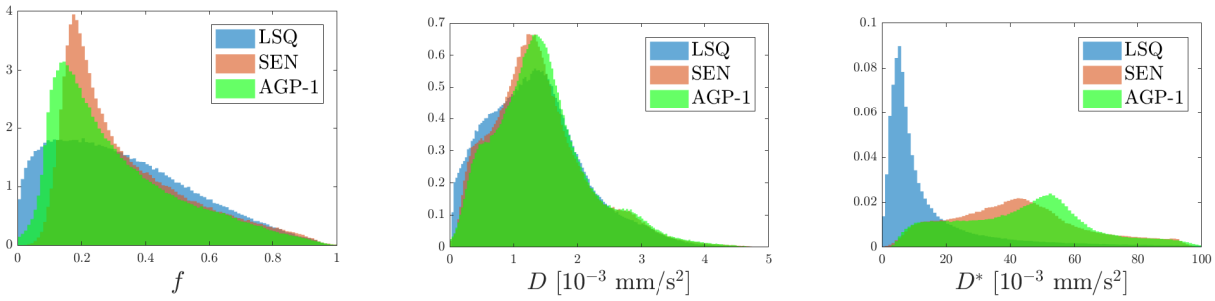


Figure 6.4: Histograms of fitted IVIM parameters by LSQ, SEN and AGP-1.

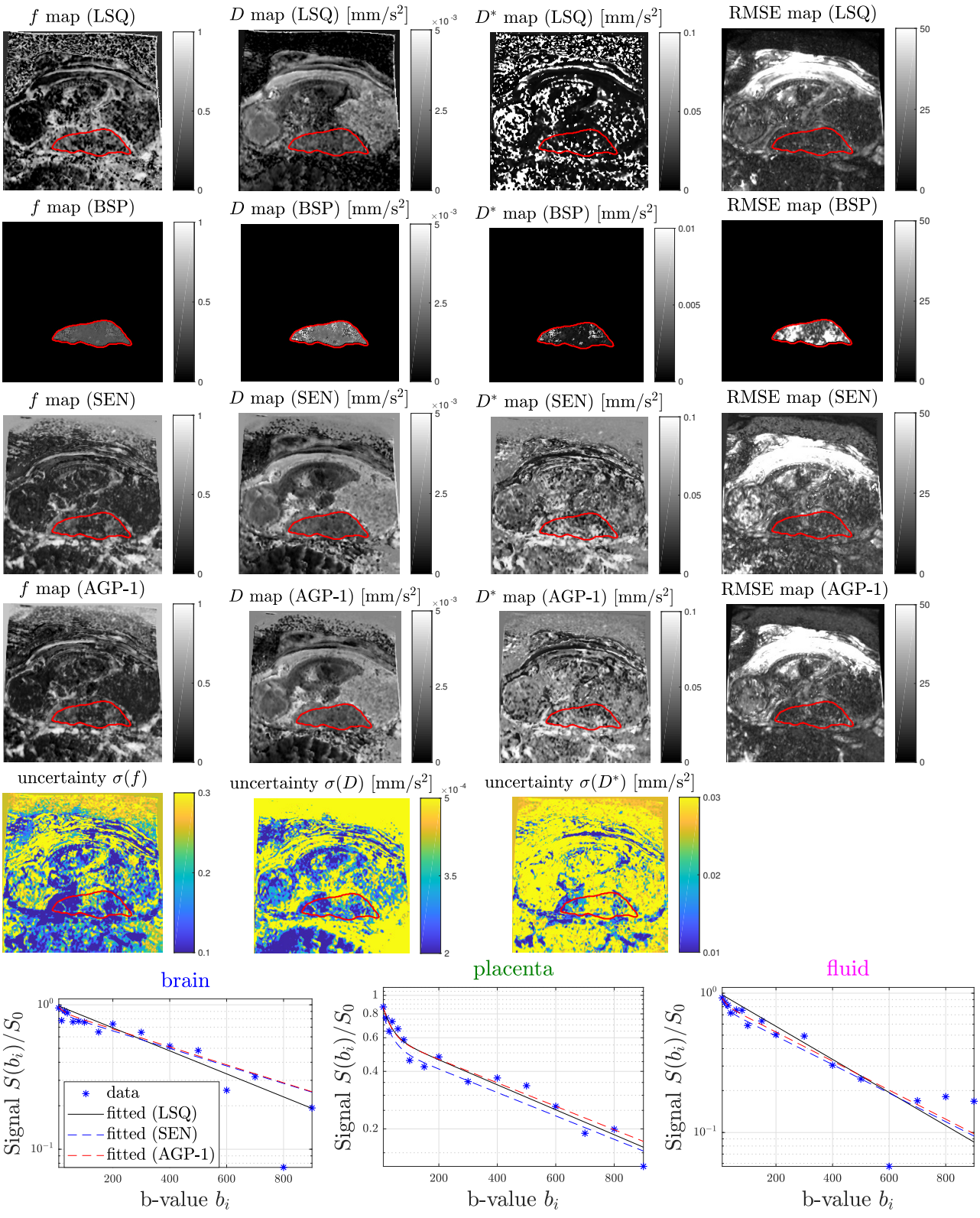


Figure 6.5: LSQ, BSP, SEN and AGP-1 parameter estimates, uncertainty maps by AGP-1, and examples of fitted signals corresponds to the example in the top row of Fig.6.1. Placenta is outlined in red.

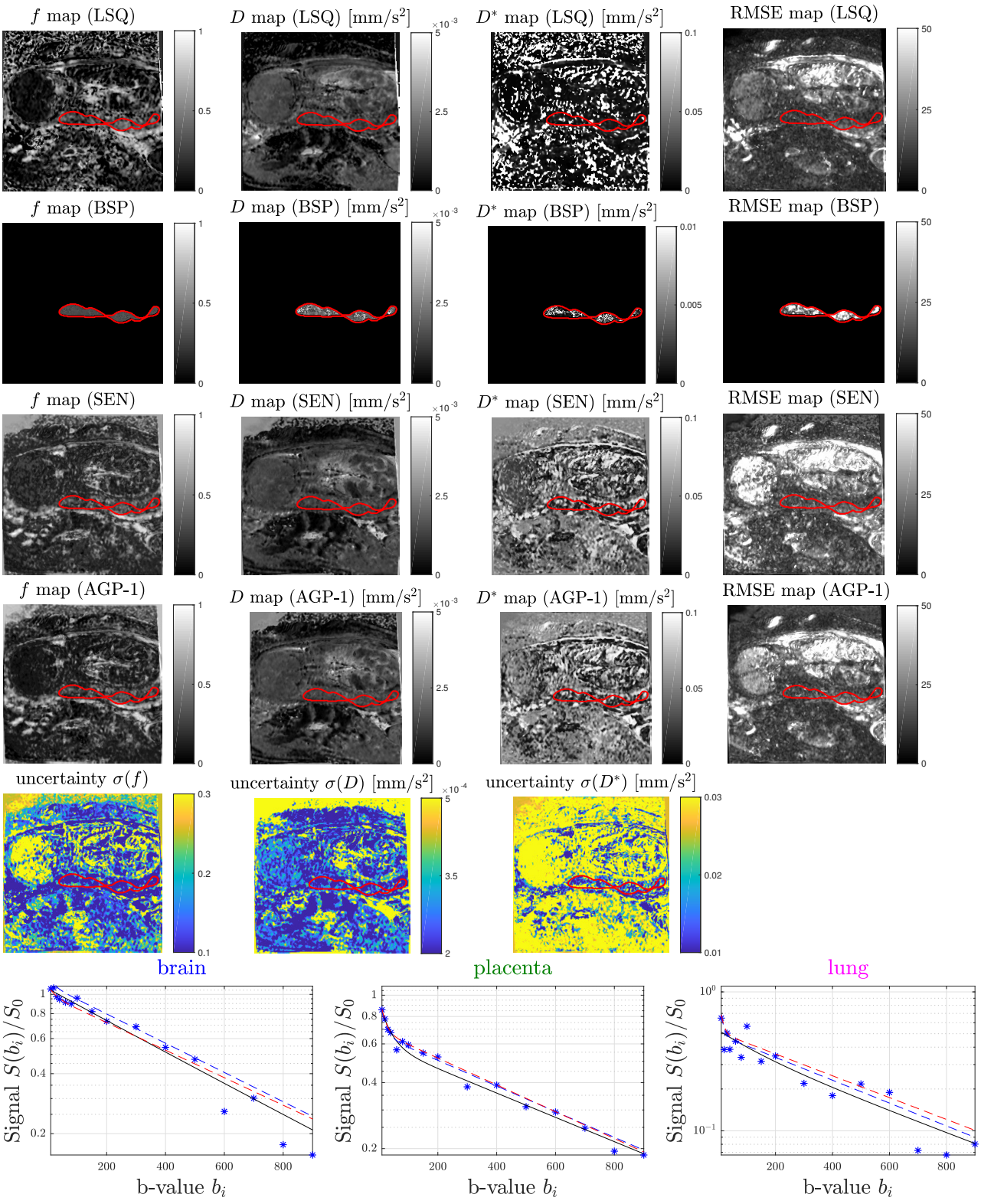


Figure 6.6: LSQ, BSP, SEN and AGP-1 parameter estimates, uncertainty maps by AGP-1, and examples of fitted signals corresponds to the example in the bottom row of Fig.6.1. Placenta is outlined in red.

Chapter 7

Discussion and Conclusion

IVIM inverse mapping for in vivo fetal MR is a challenging problem due to low SNR of images with high b-values and motion induced dephasing artefacts. To address these issues, we proposed a novel method to estimate IVIM parameters in a probabilistic way. Full posterior is modeled as mixture Gaussian with distribution parameters inferred by a multilayer perceptron with the structure defined via a coarse search over relevant hyperparameters. An implicit signal acquisition model considering non-Gaussian noise and motion induced signal loss is defined to facilitate the network training. The proposed method AGP-1 demonstrated higher estimation accuracy in terms of mean absolute error of 70% in f, 43% in D and 76% in D^* compared to the segmented least squares approach, 14% in f, 6% in D and 13% in D^* compared to SEN, on the anisotropic data with dephasing. The inferred posteriors by three variants of AGP were evaluated with the true posteriors by rejection-based Bayesian sampling, showing close approximation of the inferred posteriors by AGP-1 to the true one. Although the sampled true posteriors can be well estimated by Gaussian distribution, the estimated posteriors by AGP with two mixture components show better approximation for D^* in Fig. 5.7. Mixture density model with more mixture components can be more expressive and present as one promising direction for further investigation.

Due to lack of ground truth, the evaluation with in vivo data is hard to carry out. We use intra-subject variability as a surrogate for estimation accuracy. Higher repeatability of 51% in f, 7% in D and 60% in D^* demonstrated by AGP-1 compared to LSQ, evident for better estimation performance on the in vivo data. It is in agreement with the numerical results of the synthetic data, that AGP are less biased by dephasing artefacts. Registration does not improve the repeatability, possibly because of the erosion of the masks. Motion would induce estimation inaccuracy mainly on the boundaries between two tissues. This has been partially taken into account while delineating the placenta. Prior to model fitting, the masks are further eroded by two pixels to compensate for the movement of placenta outside the mask by the registration procedure. Therefore, the benefits of aligning images are hard to be reflected on the repeatability test with small masks. Further analysis with correlation to more reliable measurements, e.g. Doppler ultrasound, would enhance the evaluation of estimation performance in the in vivo cases. Furthermore, there exists several hypotheses about the link between the placenta perfusion and gestational age. A quadratic correlation was observed in [31]. Given the fetal MR data over a wide range of gestational weeks, investigating the correlation between the estimated parameters and gestational age can be an indicator of performance improvement, if more promising relationship can be found by applying the proposed registration and estimation algorithm.

Uncertainty map inferred by AGP-1 provides valuable information for clinical diagnosis and experimental design. For instance, while delineating region of interest, the part corrupted by motion artefacts would demonstrate high estimation uncertainty and can be ignored to avoid incorrect quantification. Moreover,

uncertainty quantification can assist with the choice of SNR level for the desired estimation precision.
The key for the successful application of AGP on the in vivo data is the closeness between the predefined implicit signal acquisition model and the measured data. The model parameters such as noise level and dephasing probability can be inferred from the data along optimizing the inference network via an adversarial training. This allows more proper forward simulation, with which the discrepancy between synthetic and measured data can be reduced. Another way is to model the signal attenuation with more descriptive biophysical model. Extending the isotropic IVIM model with tensor modeling can take diffusion and perfusion orientation into consideration. Better explanation of in vivo placenta DWI signal with anisotropic IVIM has been observed compared to isotropic IVIM in [32]. Therefore, investigation with more proper physical models appears to be a promising avenue for further work.

Appendix A

Uncertainty Quantification

Compared to the results shown in Fig. 5.8 for uncertainty quantification without dephasing simulation, here the results with dephasing simulation depicted in Fig. A.1 demonstrate similar uncertainty changes with different parameter combinations. Signal dephasing increases the estimation uncertainty overall, clearly depicted in the bottom-right plot.

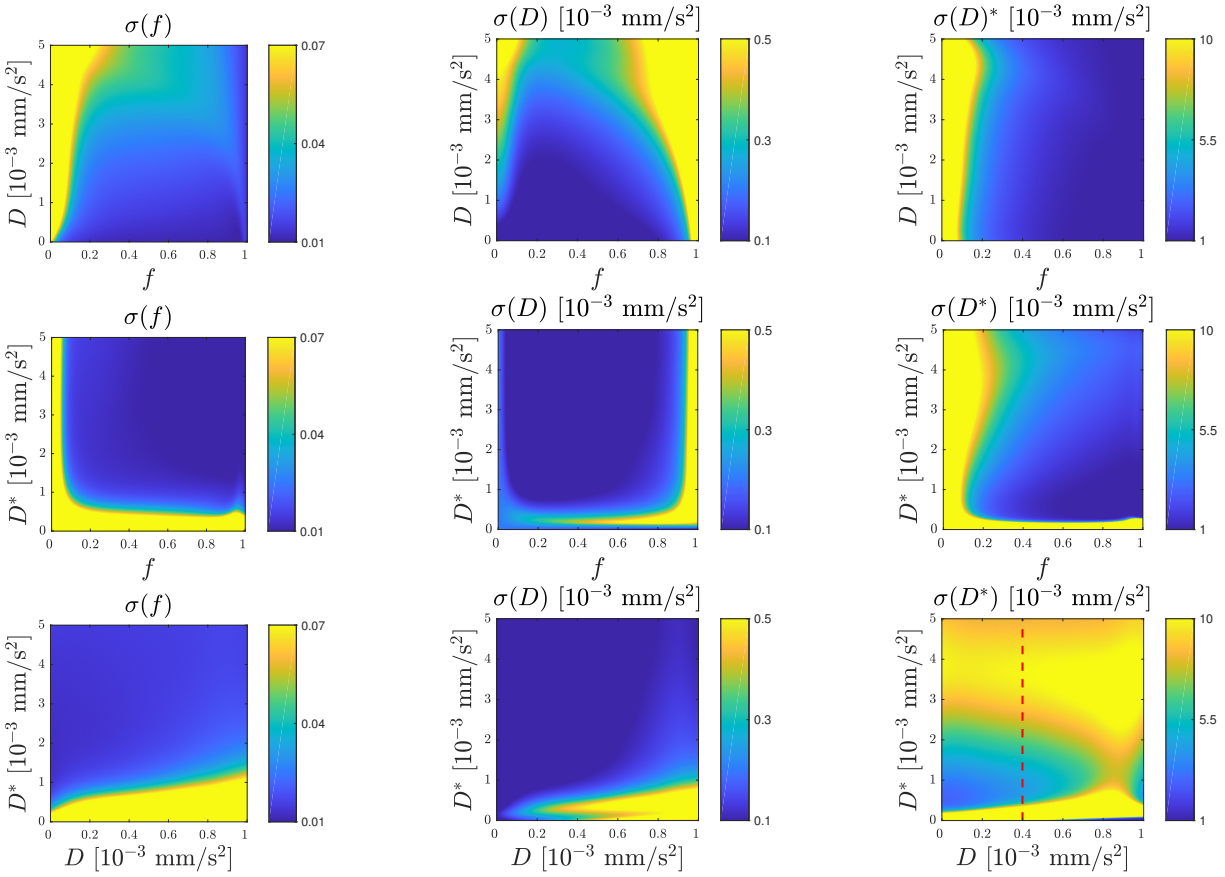


Figure A.1: Estimation uncertainty with different parameter combinations based on simulation with dephasing inferred by AGP-1. σ is the standard deviation of the predictive posterior. Top row: D^*=0.02 mm^2/s, middle: D=0.001 mm^2/s, bottom: f=0.2. S_0=1000.

Bibliography

[1] Martín Abadi, Paul Barham, Jianmin Chen, Zhifeng Chen, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Geoffrey Irving, Michael Isard, et al. Tensorflow: a system for large-scale machine learning. In *OSDI*, volume 16, pages 265–283, 2016.

[2] Marianne Alison, Gihad E Chalouhi, Gwennhael Autret, Daniel Balvay, Rokhaya Thiam, Laurent J Salomon, Charles André Cuenod, Olivier Clement, and Nathalie Siauve. Use of intravoxel incoherent motion MR imaging to assess placental perfusion in a murine model of placental insufficiency. *Investigative radiology*, 48(1):17–23, 2013.

[3] Marco Bertleff, Sebastian Domsch, Sebastian Weingärtner, Jascha Zapp, Kieran O'Brien, Markus Barth, and Lothar R Schad. Diffusion parameter mapping with the combined intravoxel incoherent motion and kurtosis model using artificial neural networks at 3 T. *NMR in Biomedicine*, 30(12):e3833, 2017.

[4] Denis Le Bihan and Robert Turner. The capillary network: a link between IVIM and classical perfusion. *Magnetic resonance in medicine*, 27(1):171–178, 1992.

[5] Sotirios Bisdas, Tong San Koh, Constantin Roder, Christian Braun, Jens Schittenhelm, Ulrike Ernemann, and Uwe Klose. Intravoxel incoherent motion diffusion-weighted MR imaging of gliomas: feasibility of the method and initial results. *Neuroradiology*, 55(10):1189–1196, 2013.

[6] Christopher M Bishop. Mixture density networks. Technical report, Neural Computing Research Group, Aston University, 1994.

[7] H Bonel, KA Frei, Luigi Raio, Matthias Meyer-Wittkopf, Luca Remonda, and Roland Wiest. Prospective navigator-echo-based real-time triggering of fetal head movement for the reduction of artifacts. *European radiology*, 18(4):822, 2008.

[8] Jon M Chia, Stefan E Fischer, Samuel A Wickline, and Christine H Lorenz. Performance of QRS detection for cardiac magnetic resonance imaging with a novel vectorcardiographic triggering method. *Journal of Magnetic Resonance Imaging*, 12(5):678–688, 2000.

[9] Thomas E Conturo, Robert C McKinstry, Erbil Akbudak, and Bruce H Robinson. Encoding of anisotropic diffusion with tetrahedral gradients: a general mathematical diffusion formalism and experimental results. *Magnetic Resonance in Medicine*, 35(3):399–412, 1996.

[10] Katalin Csilléry, Michael GB Blum, Oscar E Gaggiotti, and Olivier François. Approximate Bayesian computation (ABC) in practice. *Trends in ecology & evolution*, 25(7):410–418, 2010.

[11] Richard L Ehman, MT McNamara, M Pallack, H Hricak, and CB Higgins. Magnetic resonance imaging with respiratory gating: techniques and advantages. *American journal of Roentgenology*, 143(6):1175–1182, 1984.

[12] Tim Finkenstaedt, Markus Klarhoefer, Christian Eberhardt, Anton S Becker, Gustav Andreisek, Andreas Boss, and Cristina Rossi. The IVIM signal in the healthy cerebral gray matter: a play of spherical and non-spherical components. *Neuroimage*, 152:340–347, 2017.

[13] John C Haselgrove and James R Moore. Correction for distortion of echo-planar images used to calculate the apparent diffusion coefficient. *Magnetic Resonance in Medicine*, 36(6):960–964, 1996.

[14] Wyke Huizinga, Dirk HJ Poot, J-M Guyader, R Klaassen, Bram F Coolen, Matthijs van Kranenburg, RJM Van Geuns, André Uitterdijk, Mathias Polfliet, Jef Vandemeulebroucke, et al. PCA-based groupwise image registration for quantitative MRI. *Medical image analysis*, 29:65–78, 2016.

[15] András Jakab, Ruth Tuura, Raimund Kottke, Christian J Kellenberger, and Ianina Scheer. Intra-voxel incoherent motion MRI of the living human foetus: technique and test–retest repeatability. *European radiology experimental*, 1(1):26, 2017.

[16] András Jakab, Ruth L Tuura, Raimund Kottke, Nicole Ochsenbein-Kölble, Giancarlo Natalucci, Thi Dao Nguyen, Christian Kellenberger, and Ianina Scheer. Microvascular perfusion of the placenta, developing fetal liver, and lungs assessed with intravoxel incoherent motion imaging. *Journal of Magnetic Resonance Imaging*, 48(1):214–225, 2018.

[17] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

[18] Sila Kurugol, Moti Freiman, Onur Afacan, Liran Domachevsky, Jeannette M Perez-Rossello, Michael J Callahan, and Simon K Warfield. Motion compensated abdominal diffusion weighted MRI by simultaneous image registration and model estimation (SIR-ME). In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 501–509. Springer, 2015.

[19] P Lanzer, C Barta, EH Botvinick, HU Wiesendanger, G Modin, and CB Higgins. ECG-synchronized cardiac MR imaging: method and evaluation. *Radiology*, 155(3):681–686, 1985.

[20] Denis Le Bihan. What can we see with IVIM MRI? *NeuroImage*, 2017.

[21] Denis Le Bihan, Eric Breton, Denis Lallemand, ML Aubin, J Vignaud, and M Laval-Jeantet. Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology*, 168(2):497–505, 1988.

[22] CE Lewis, FS Prato, DJ Drost, and RL Nicholson. Comparison of respiratory triggering and gating techniques for the removal of respiratory artifacts in MR imaging. *Radiology*, 160(3):803–810, 1986.

[23] Alain Luciani, Alexandre Vignaud, Madeleine Cavet, Jeanne Tran Van Nhieu, Ariane Mallat, Lucile Ruel, Alexis Laurent, Jean-François Deux, Pierre Brugieres, and Alain Rahmouni. Liver cirrhosis: intravoxel incoherent motion MR imagingpilot study. *Radiology*, 249(3):891–899, 2008.

[24] C Malamateniou, SJ Malik, SJ Counsell, JM Allsop, AK McGuinness, T Hayat, K Broadhouse, RG Nunes, AM Ederies, JV Hajnal, et al. Motion-compensation techniques in neonatal and fetal MR imaging. *American Journal of Neuroradiology*, 34(6):1124–1136, 2013.

[25] Klaus-Dietmar Merboldt, Wolfgang Hänicke, and Jens Frahm. Diffusion imaging using stimulated echoes. *Magnetic resonance in medicine*, 19(2):233–239, 1991.

[26] David A Nix and Andreas S Weigend. Estimating the mean and variance of the target probability distribution. In *Neural Networks, 1994. IEEE World Congress on Computational Intelligence., 1994 IEEE International Conference On*, volume 1, pages 55–60. IEEE, 1994.

[27] Mike Notohamiprodjo, Hersh Chandarana, Artem Mikheev, Henry Rusinek, John Grinstead, Thorsten Feiweier, José G Raya, Vivian S Lee, and Eric E Sigmund. Combined intravoxel incoherent motion and diffusion tensor imaging of renal diffusion and flow anisotropy. *Magnetic resonance in medicine*, 73(4):1526–1532, 2015.

[28] Matthew R Orton, David J Collins, Dow-Mu Koh, and Martin O Leach. Improved intravoxel incoherent motion analysis of diffusion weighted imaging by data driven Bayesian modeling. *Magnetic resonance in medicine*, 71(1):411–420, 2014.

[29] Gustavo Kunde Rohde, AS Barnett, PJ Basser, S Marenco, and C Pierpaoli. Comprehensive approach for correction of motion and distortion in diffusion-weighted MRI. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, 51(1):103–114, 2004.

[30] Hiroshi Shinmoto, Koichi Oshio, Akihiro Tanimoto, Nobuya Higuchi, Shigeo Okuda, Sachio Kuribayashi, and Robert V Mulkern. Biexponential apparent diffusion coefficients in prostate cancer. *Magnetic resonance imaging*, 27(3):355–359, 2009.

[31] Nathalie Siauve, Pierre Humbert Hayot, Benjamin Deloison, Gihad E. Chalouhi, Marianne Alison, Daniel Balvay, Laurence Bussires, Olivier Clment, and Laurent J. Salomon. Assessment of human placental perfusion by intravoxel incoherent motion MR imaging. *The Journal of Maternal-Fetal & Neonatal Medicine*, 0(0):1–8, 2017. PMID: 28974131.

[32] Paddy J Slator, Jana Hutter, Laura McCabe, Ana Dos Santos Gomes, Anthony N Price, Eleftheria Panagiotaki, Mary A Rutherford, Joseph V Hajnal, and Daniel C Alexander. Placenta microstructure and microcirculation imaging with diffusion MRI. *Magnetic resonance in medicine*, 80(2):756–766, 2018.

[33] Georg R Spinner, Constantin Von Deuster, Kerem C Tezcan, Christian T Stoeck, and Sebastian Kozerke. Bayesian intravoxel incoherent motion parameter mapping in the human heart. *Journal of Cardiovascular Magnetic Resonance*, 19(1):85, 2017.

[34] Edward O Stejskal and John E Tanner. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, 42(1):288–292, 1965.

[35] Christian et al. Stoeck. Analysis of cardiac motion induced error for in-vivo cardiac DTI in different heart phases. a comparison of second-order motion compensated SE versus STEAM. *ISMRM*, 2018.

[36] Petre Stoica and Arye Nehorai. MUSIC, maximum likelihood, and Cramer-rao bound. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, 37(5):720–741, 1989.

[37] Joshua FP van Amerom, David FA Lloyd, Anthony N Price, Maria Kuklisova Murgasova, Paul Aljabar, Shaihan J Malik, Maelene Lohezic, Mary A Rutherford, Kuberan Pushparajah, Reza Razavi, et al. Fetal

cardiac cine imaging using highly accelerated dynamic MRI with retrospective motion correction and outlier rejection. *Magnetic resonance in medicine*, 79(1):327–338, 2018.

[38] Valery Vishnevskiy, Tobias Gass, Gabor Szekely, Christine Tanner, and Orcun Goksel. Isotropic total variation regularization of displacements in parametric image registration. *IEEE transactions on medical imaging*, 36(2):385–395, 2017.

[39] Ichiro Yamada, Winn Aung, Yoshiro Himeno, Tsuneaki Nakagawa, and Hitoshi Shibuya. Diffusion coefficients in abdominal organs and hepatic lesions: evaluation with intravoxel incoherent motion echo-planar MR imaging. *Radiology*, 210(3):617–623, 1999.

[40] Qinwei Zhang, Yi-Xiang Wang, Heather Ting Ma, and Jing Yuan. Cramér-Rao bound for intravoxel incoherent motion diffusion weighted imaging fitting. In *Engineering in Medicine and Biology Society (EMBC), 2013 35th Annual International Conference of the IEEE*, pages 511–514, 2013.